

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

June 18, 2003, 12:43:22 ; Search time 36 Seconds  
(without alignments)  
2878.936 Million cell updates/sec

09/647, 965  
Search for  
#2

Title: US-09-647-965-9  
Perfect score: 2731  
Sequence: 1 MALAPERARAVLFGFWLIG.....SANSLYDIECFLEMELOPA 503

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

1: SPTRMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhc:\*  
8: sp\_mammal:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rviro:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1116	40.9	210 4 Q9UE79	Q9UE79 homo sapien
2	778	28.5	491 13 Q90ZD4	Q90ZD4 gallus galli
3	603	22.1	128 4 Q9HB64	Q9HB64 homo sapien
4	400	16.1	459 13 Q57578	Q57578 xenopus lae
5	430	15.7	467 6 Q9N136	Q9N136 civis aries
6	428	15.7	467 11 Q91VD0	Q91VD0 mus musculu
7	420	15.4	452 4 Q8WNC4	Q8WNC4 sus scrofa
8	364.5	13.3	445 4 Q96GJ3	Q96GJ3 homo sapien
9	364	13.3	445 13 Q98TX7	Q98TX7 gallus galli
10	362	13.3	409 13 Q98TX6	Q98TX6 gallus galli
11	344	12.6	440 4 Q99419	Q99419 homo sapien
12	321	11.8	416 13 Q90W10	Q90W10 gallus galli
13	281.5	10.3	330 11 Q90ZL7	Q90ZL7 mus musculu
14	194	7.1	215 11 Q924T6	Q924T6 mus musculu
15	183	6.7	1334 16 Q9RKR9	Q9RKR9 streptomyce
16	177.5	6.5	296 13 Q98TVA1	Q98TVA1 fungu rubrip

17	176	6.4	116 13 Q90584	Q90584 gallus galli
18	173.5	6.4	322 6 Q9BG75	Q9BG75 ovis aries
19	169.5	6.2	349 6 Q9N137	Q9N137 ovis aries
20	168.5	6.2	297 13 Q9PW19	Q9PW19 paratichthy
21	166	6.1	349 11 Q923E9	Q923E9 mus musculu
22	166	6.1	349 11 Q8R4E0	Q8R4E0 sigmodon hi
23	165.5	6.1	469 12 Q69270	Q69270 equine herp
24	165	6.0	3179 12 Q8V2A4	Q8V2A4 human herp
25	164	6.0	315 13 Q91838	Q91838 coturnix co
26	163	6.0	91 6 Q95290	Q95290 sus scrofa
27	163	6.0	880 11 Q971A0	Q971A0 cavia porce
28	162.5	6.0	385 12 Q41935	Q41935 murid herpe
29	162	5.9	98 13 Q90W55	Q90W55 ctenopharyn
30	162	5.9	1660 5 Q9NKN0	Q9NKN0 leishmania
31	161.5	5.9	551 4 Q9Y613	Q9Y613 homo sapien
32	161	5.9	889 16 Q9F2N5	Q9F2N5 streptomyce
33	160.5	5.9	1400 5 Q9NA90	Q9NA90 caenorhabdi
34	159	5.8	576 16 Q9XA04	Q9XA04 streptomyce
35	158.5	5.8	694 6 Q9GKY7	Q9GKY7 oryctolagus
36	158.5	5.8	839 10 Q9SN46	Q9SN46 arabidopsis
37	158.5	5.8	1463 16 Q9ADP6	Q9ADP6 streptomyce
38	157.5	5.8	3247 12 Q65553	Q65553 bovine herp
39	156	5.7	381 5 Q94399	Q94399 caenorhabdi
40	156	5.7	552 4 Q9BR39	Q9BR39 homo sapien
41	155	5.7	576 4 Q9HA18	Q9HA18 homo sapien
42	155	5.7	678 11 Q80ZT8	Q80ZT8 mus musculu
43	155	5.7	801 5 Q23635	Q23635 caenorhabdi
44	154	5.6	696 11 Q9ET79	Q9ET79 mus musculu
45	154	5.6	1691 11 Q9ESQ2	Q9ESQ2 mus musculu

## ALIGNMENTS

## RESULT 1

Q9UE79 Q9UE79 PRELIMINARY; PRT; 210 AA.  
AC Q9UE79;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Putative interferon regulatory factor 7C.2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN, AND BREAST;  
RA MEDLINE=97459673; PubMed=9315633;  
RA Zhang L.; Pagano J.S.;  
RT "IRF-7", a new interferon regulatory factor associated with Epstein-  
RT Barr virus latency."  
RL Mol. Cell. Biol. 17:5748-5757(1997).  
DR EMBL: U53832; AAB80691.1; -;  
SQ SEQUENCE 210 AA; 23592 MW; 45895671CEAA18F5 CRC64;

Query Match 40.9%; Score 1116; DB 4; Length 210;

Best Local Similarity 99.0%; Pred. No. 9.5e-70;

Matches 208; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	294	MYKGRVYLVKQVGHPSCTFLYGPDPDAVRATDPOQVAFPSAPALPPQKOLRYTEELLRHV	353
Db	1	MYKGRVYLVKQVGHPSCTFLYGPDPDAVRATDPOQVAFPSAPALPPQKOLRYTEELLRHV	60
Qy	354	APGLHLELRGPQIMARRMKCKYWEVGGPGSASPSTPACLLPNCDDTPIFDPRVFFOE	413
Db	61	APGLHLELRGPQIMARRMKCKYWEVGGPGSASPSTPACLLPNCDDTPIFDPRVFFOE	120
Qy	414	LVEFRARORGRSPRYTYIGFGODLSAGRPKESLVLVLEPWLCHVHLEGTORREGVSSL	473
Db	121	LVEFRARORGRSPRYTYIGFGODLSAGRPKESLVLVLEPWLCHVHLEGTORREGVSSL	180

OY 474 DSSDLCLSSANSIYDIECFMELEOPA 503  
 DB 181 DSSSLCLSSANSIYDIECFMELEOPA 210

## RESULT 2

OY902D4 PRELIMINARY; PRT; 491 AA.  
 AC O902D4;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Interferon regulatory factor-3.  
 GN IRF-3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA May D.L., Grant C.E., Deeley R.G.;  
 RT "Cloning and Promoter Analysis of the Chicken Interferon Regulatory  
 Factor-3 Gene."  
 DT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF268079; AAK58583.1; -  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR ProDom; PD002355; IRF.1.  
 DR PROSITE; PS00601; IRF; UNKNOWN.1.  
 SQ SEQUENCE 491 AA; 5440 MW; 90E0C80A2624B1B4 CRC64;

Query Match 28.5%; Score 778; DB 13; Length 491;  
 Best Local Similarity 38.5%; Pred. No. 5.5e-46;  
 Matches 208; Conservative 59; Mismatches 181; Indels 92; Gaps 17;

OY 2 ALAPRAAPVLFGEMLGEISSGCEYGLQWLDARTCFRVWPKFARKDSEADARIFK 61  
 DB 3 ALDSGDAQKLFKFGFWLNAVSSGUYRGICWIDPRRIIPRIWKNARKDVTSSDVEIFK 62  
 OY 62 AMAVARGRPSSRGCGPPPEAETARACWKTNPFCALSTRFVWLRNDSGDPADPHKV 121  
 DB 63 AWAASGRY-----EGNADPAKMTNFCALSTRFVWLRNDSGDPADPHKV 111  
 OY 122 YALSRELQWREGPTDQTEAABAAVPPPO--GPPFLAHTAGH-- 167  
 DB 112 YAVA-----SGVPNRSGSGGFPVAGALQOQPOLLNHDLAENTPT 152  
 OY 168 -APGFLPAPADGKDL-LIQAVQSGCLADHLITASWGADPVPTKAPGEG--QEGPLTG- 222  
 DB 153 DSTEGLVAAAALQVLDLQSVLQHCNISAL-----GSGPTLMAHTGDALPEDALLPQ 207  
 OY 223 -ACAGPGPLPAGELXGW-AVETTPSPGPP-----AALTGGAAPESHQNEP 269  
 DB 208 DGCLPQP-----QFQDWRQLEPRLLLGNQPLTGGCGGQDAGALPVSECAIPAPSPAE 262  
 OY 270 YL-----SPSPACTAIVQEPSPGALDVTIYKGRITVLOKVGHPSTCTIYGPDPAPVATD 325  
 DB 263 LIFQGANRPPPPADIGLPPPLDITTYRGRKMTYQEVDSRCVLAQPLDPVAV--AE 320  
 OY 326 PQQVAPSPFAPLIPDOKLRYTELLRHVAPGLHLERGQVLMARWAKCKYWEV----- 380  
 DB 321 QRLVFPSPASLPDRQRYTENLE--VAGRLRQAGQLATRLKCKKCVFMAISQLE 378  
 OY 381 GPPPGASSTPACLLPNCOTPIPDFRFVPELVFPRARQSGSPRTIYIGFRODLSA 440  
 DB 379 GGP-----PLNLHRDQETITDFRVCTELDFRGRERSPDFTIFLCFQCSS 431  
 OY 441 GPPKESLIVLKEPILCRVHLEGTQREGVSSLDSSDDLCLSSANSIYDIECFMELE 500  
 DB 432 TKPKSKLIVLVQFCGYWYEQVQRGASSLNGNSVLQSDSFNLFELIEQYHMQTD 491

## RESULT 3

OY9HB64 PRELIMINARY; PRT; 128 AA.  
 AC O9HB64;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Interferon regulatory factor-7H (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20493557; PubMed=10924517;  
 RA Lu R., Au W.-C., Yeow M.-S., Hageman N., Pitha P.M.;  
 RT "Regulation of the Promoter Activity of Interferon Regulatory Factor-7  
 gene Activation by Interferon and Silencing by Hypermethylation."  
 RL J. Biol. Chem. 275:31805-31812 (2000).  
 DR EMBL; AF277159; AAG30003.1; -  
 DR HSSP; P23906; 2IRF.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERREGCT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 FT NON TER 128  
 SQ SEQUENCE 128 AA; 14363 MW; 09B5B6F3C11BB15 CRC64;

Query Match 22.1%; Score 603; DB 4; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-34;  
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 RAARVLFGEMLGEISSGCEYGLQWLDARTCFRVWPKFARKDSEADARIFKAWA 66  
 DB 20 RAARVLFGEMLGEISSGCEYGLQWLDARTCFRVWPKFARKDSEADARIFKAWA 79  
 OY 67 RGRPPSSRGCGPPPEAETARACWKTNPFCALSTRFVWLRNDSGDP 115  
 DB 80 RGRPPSSRGCGPPPEAETARACWKTNPFCALSTRFVWLRNDSGDP 128

## RESULT 4

OY57578 PRELIMINARY; PRT; 459 AA.  
 AC O57578;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE XIRF-6 protein.  
 GN XIRF-6.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98086218; PubMed=9426249;  
 RA Hatada S., Kinoshita M., Takahashi S., Nishihara R., Sakumoto H.,  
 RA Fukui A., Noda M., Asashima M.;  
 RT "An Interferon Regulatory Factor-Related Gene (XIRF-6) is Expressed in  
 the Posterior Mesoderm during the Early Development of Xenopus  
 laevis."  
 RL Gene 203:183-186 (1997).  
 DR EMBL; D86492; BAA24349.1; -  
 DR HSSP; P23906; 2IRF.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERREGCT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.

DR PROSITE; PS00601; IRF; 1.  
SQ SEQUENCE 459 AA; 52475 MW; 306DA77C6586871C CRC64;

Query Match 16.1%; Score 440; DB 13; Length 459;  
Best Local Similarity 27.1%; Pred. No. 1,1e-22;  
Matches 140; Conservative 74; Mismatches 197; Indels 106; Gaps 19;

QY 1 MALAPRAARVLFGEWLTGELISGCEYGLQWLDARTCFRVWPKFAR-KDISEADARI 59  
D 1 MAHPR-----RVRLKPLVAQVDSGMYPGILINREAKRFQIPWKATRISSPOGEENIT 56  
QY 60 FKAAVARGWPPSSRGCGPPPEAEIARAGWKTFRCAISTRFVMLRDNAGD-PADP 118  
D 57 FKAAVETGKYQEG-----ADEDDPAKWKQALCALINKSRFEFLMTDGTKEVPMNP 107  
QY 119 HKYVALSRELCEWREGPTDQTEAARPAVPPGCGPPFLATHAGLQAPGLPAPAGD 178  
D 108 VKIY---EVC-----DIPQSG-----SINPGSTGSVPMD 135  
QY 179 KGDLLLAQAVQSCILADHLITASWGADVPPTKAP-----GEGEGLEPLTGACAGGFLPA 232  
D 136 DDPEADELNGS-----QNHVPISFPNCININDSPIGSSITSGCT----- 176  
QY 233 GELYGMAVETTPSPGPAPAL--TTGEAAAPESPHQAEPLYSPSPSACTAVQEPSPALD 290  
D 177 -----PEQWPKTEPEQEMEVPPPTSGPADFFSSP---EMWISLSPMT-----DLE 217  
QY 291 VTIMYKRTVLYQY-VGHP-SCTEFLYGPPDPVAVATD-----PQVAFPSPAELPDQK 342  
D 218 IQFYRKKEGQITVNSPQGCRLFYDGLGMPNOEELFGPTLEQVRPGTBOIVNEK 277  
QY 343 LRYTEELLRHVAPGLHLRLGPPQLMARMGCKKYWEVGGPPSGASSTPACLLPRNCDT 402  
D 278 KLFTSRLLDVMRDLGLIEVSGHAIYAIRLCQCKYV---SGCSPSITTN-FIERQKV 333  
QY 403 PTFDFRVFOELVEFRARQRRGSPRTIYLGFGODLSAGRPKESLYLVKLEPMLCRVHL 462  
D 334 KLCEVFETFLDLISHQGIITTKQPPYIYLGFGEMWDGKYEKRLIYQIPIVAMMI 393  
QY 463 EGTQREGVSLDSSDLCLSSANSLYDIECFMEL 499  
D 394 EMFTGSTRSFGSGIRLQIIPD-IKDNIVSHLKL 429

RESULT 5

Q9N136 PRELIMINARY; PRT; 467 AA.

AC Q9N136; (TREMBlurel. 15, Created)  
DT 01-OCT-2000 (TREMBlurel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlurel. 19, Last annotation update)  
DE 01-DEC-2001 (TREMBlurel. 19, Last annotation update)  
OS Interferon regulatory factor 6.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE=15 DAY PREGNANT UTERUS;  
RC Choi Y., Spencer T.E., Bazer F.W.;  
RT "Cloning and Analysis of Ovine IRF-6";  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF228446; AF34782.1; -.  
DR HSSP; P23906; 21RF.  
DR InterPro; IPR001346; IRF.  
DR Pfam; PF00605; IRF; 1.  
DR PRINTS; PR00267; INTERFERECT.  
DR ProDom; PD002355; IRF; 1.  
DR SMART; SM00348; IRF; 1.  
DR PROSITE; PS00601; IRF; 1.  
SQ SEQUENCE 467 AA; 52970 MW; 21E04F749844D88F CRC64;

Query Match 15.7%; Score 430; DB 6; Length 467;  
Best Local Similarity 27.3%; Pred. No. 5,7e-22;  
Matches 141; Conservative 72; Mismatches 209; Indels 94; Gaps 20;

QY 1 MALAPRAARVLFGEWLTGELISGCEYGLQWLDARTCFRVWPKFAR-KDISEADARI 59  
D 1 MAHPR-----RVRLKPLVAQVDSGLYPGLIWHRDSKRFQIPWKATRISSPOGEENIT 56  
QY 60 FKAAVARGWPPSSRGCGPPPEAEIARAGWKTFRCAISTRFVMLRDNAGD-PADP 118  
D 57 FKAAVETGKYQEG-----QGVDDPDPK-----WKAQRLCALINKSRERNIMTGTKEVPMNP 107  
QY 119 HKYVALSRELCEWREGPTDQTEAARPAVPPGCGPPFLATHAGLQAPGLP----- 173  
D 108 VKIY---QVC-----DIPQSGSIINP-----GSTGAPMDK 138  
QY 174 --APAGDKDLLLAQAVQSCILADHLITASWGADVPPTKAPGEGEGLEPLTGACAGGFLP 231  
D 139 NDVDEBDEBDELQSGHVPIDQTFPELNINGSPI--APGS-----VGNCSVNCSP 188  
QY 232 AGELYGMAVETTPSPGPAPALTTGEAAAPESPHQAEPLYSPSPSACTAVQEPSPALDV 291  
D 189 EAV---W-----PTPELEMEVPPQ-APIQFYSPEIMISLSPMT-----DLDI 228  
QY 292 TIMYKRTVLYQY-VGHP-SCTEFLYGPPDP-----AVRATDQVAFPSPAELPDQK 343  
D 229 KFYRGKEVGQITVNSPQGCRLFYDGLGMPNOEELFGVSLQVKFPPEHITNBK 288  
QY 344 RYTEELLRHVAPGLHLRLGPPQLMARMGCKKYWEVGGPPSGASSTPACLLPRNCDT 403  
D 289 LFTSKLLDVMRDLGLIEVSGHAIYAIRLCQCKYWGCPAFLVAPN---LIERQKV 344  
QY 404 IFDFRVFOELVEFRARQRRGSPRTIYLGFGODLSAGRPKESLYLVKLEPMLCRVHL 463  
D 345 LFLFTFLDGLIHHGQLEKQPPREIYLCFGEWWDGKYEKRLIYQIPIVAAKITE 404  
QY 464 GGTQREGVSLDSSDLCLSSANSLYDIECFMEL 499  
D 405 MFGDFTSRSGSVRLQISTPD-IKDNIVAQKQL 439

RESULT 6

Q91VD0 PRELIMINARY; PRT; 467 AA.

AC Q91VD0; (TREMBlurel. 19, Created)  
DT 01-DEC-2001 (TREMBlurel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlurel. 21, Last annotation update)  
DE BM282D4.4 (Interferon regulatory factor 6).  
GN IRF6.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tromans A.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL365322; CAC42184.1; -.  
DR EMBL; BC008515; AA08515.1; -.  
DR MGD; MG1:1859211; 11F6.  
DR InterPro; IPR001346; IRF.  
DR Pfam; PF00605; IRF; 1.  
DR ProDom; PD002355; IRF; 1.  
DR PROSITE; PS00601; IRF; 1.  
SQ SEQUENCE 467 AA; 53110 MW; 1C564BC8D79C5259 CRC64;

Query Match 15.7%; Score 428; DB 11; Length 467;  
Best Local Similarity 27.5%; Pred. No. 7,8e-22;

Matches 141; Conservative 72; Mismatches 212; Indels 88; Gaps 17;

```

OY 1 MALABERAPRLFEWMLLGEISSGCGYBGLQWLDARCTCFRVPWGHFAR-KDLSADARI 59
DB 1 MALHRR-----RVRLKPMVLAVQVSGLYPGLIMLHRDSKRFOIPWGHARHSPQOEENIT 56
OY 60 FKAAVAVGRWPPSSRGCGPPPEAETARAGKTNFRCAIARSTRFRVWLNDNSGD-PADP 118
DB 57 FKAAVAVETGKY--QEGVDDPDPK-----WKQRLCALNKSREFNLMYDGTKEVPMNP 107
OY 119 HKVYALSELCKMBEGPTDQTEAEPAAVPPGPGPFLAHTHAGLQAPGAPLAPAGD 178
DB 108 VKIY-----QVC-----DIPQPG-----SVINPGSTGSAPWD 135
OY 179 KGDLLQAVQGSCLADHLITASWGADEVPTKAPGEGGLPL-----TGACAGGGLPAGE 234
DB 136 EKDNVDDEDE---EDELQSQHNVPIQDTFFPLNINGSPIAPASVNCISGNCSP--- 188
OY 235 LYGMVAVETTPSPGPAPALTGGEAAPSPHQAEPYLSPPSPACTAVOEPSGALDVTIM 294
DB 189 -----ESVMPKTEPLEMEVPAPIQPFYSSPELMISSLPMT-----DDIKFO 231
OY 295 YKGRVTLOKV-VGHP-SCTFLYGPDP-----AVRATDPOQVAFPSPAELPDOKOLRYT 346
DB 232 YRGEYQGTMTVSNPQCGRLFYGDIGPMPDQELFGPVSLBQVKPGEHITNEKOLFT 291
OY 347 BELLRHVAFGHLERGPOLMARBMGCKYVWEVGGPGSASPTPACLLPRNCDTPIFD 406
DB 292 SKLLVMPORGILLVESHAIYAIRLCQCKYVWSGPCAPSLAPN---LIEROKKVKLFC 347
OY 407 FRVFPQELVEFRARQSGSPRYTILYFGODLSAGRPKESLVLYKLEPMLCRVHLESTQ 466
DB 348 LETFSELIAHQKQIEKOPPEIYLCGEBWPDCKPBERKILIVQVIVPVARMIYEMFS 407
OY 467 REGVSLDSSDLDCLSSANSLYDIECFMEL 499
DB 408 GDFTRSPDSGSVRLQISTPD-IKDNIVAQLKOL 439

RESULT 7
OY 08BNQ4 PRELIMINARY; PRT; 467 AA.
AC 08BNQ4;
DB 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Interferon regulatory factor 6.
GN IRF6.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
NC NCBI_Taxid=9623;
RP SEQUENCE FROM N.A.
RC TISSUE=ILEUM;
RA Father C.R., Raney N.E., Ernst C.W.;
RT "Characterization of the porcine IRF6 gene: cDNA cloning, expression
RT analysis and chromosomal localization."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327368; AAL37429.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERNEGCT.
DR ProDom; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF; UNKNOWN.1.
SQ SEQUENCE 467 AA; 53047 MW; 4AB757DA8013A3C2 CRC64;

Query Match 15.4%; Score 420; DB 6; Length 467;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 138; Conservative 71; Mismatches 210; Indels 84; Gaps 19;
OY 11 RVLFGEWMLLGEISSGCGYBGLQWLDARCTCFRVPWGHFAR-KDLSADARIFKAAVAVGR 69

```

```

DB 7 RVRLKPMVLAVQVSGLYPGLIMLHRDSKRFOIPWKHATRRSPQOEENITFKAAVAVETGK 66
OY 70 WPPSSRGCGPPPEAETARAGKTNFRCAIARSTRFRVWLNDNSGD-PADPHKYALREL 128
DB 67 Y--QEGVDDPDPK-----WKQRLCALNKSREFNLMYDGTKEVPMNPVKIY---QV 113
OY 129 CMREGPTDQTEAEPAAVPPGPGPFLAHTHAGLQAPGAPLAPAGDGLLLQAVQ 188
DB 114 C-----DIPQPG-----SVNPGSTGSAPMEKD---NDVD 142
OY 189 QSCLDHLITASWGADEVPTKAPGEGGLPLGA-----CAGGGLPAGELYGMVAVETTP 244
DB 143 DEDEDE-LDQSQHNVPIQDTFFPLNINGSPIAPASVNCISGNCSP---W----- 192
OY 245 SPGPAPALTGGEAAPSPHQAEPYLSPPSPACTAVOEPSGALDVTIMYKGRVTLOKV 304
DB 193 -PKAEPLEMEVQ--APIQPFYSSPELMISSLPMT-----DLDIRFOYRGEYQGTM 241
OY 305 -VGHP-SCTFLYGPDP-----AVRATDPOQVAFPSPAELPDOKOLRYTELLRHVAPG 356
DB 242 TVSNPQCGRLFYGDIGPMPDQELFGPVSLBQVKPGEHITNEKOLFTSKLLDWDNRG 301
OY 357 LHLELRGPOLMARBMGCKYVWEVGGPGSASPTPACLLPRNCDTPIFDPRVFFQELVE 416
DB 302 LIEVSHAIYAIRLCQCKYVWSGPCAPSLAVAPN---LIEROKKVKLFCLEFTLSDLIA 357
OY 417 FRANQRRGSPRYTILYFGODLSAGRPKESLVLYKLEPMLCRVHLESTQREGVSLDSS 476
DB 358 HQKQIEROPPEIYLCGEBWPDCKPBERKILIVQVIVPVARMIYEMFSGDFTFRSPDSG 417
OY 477 DLDCLSSANSLYDIECFMEL 499
DB 418 SVRLQISTPD-IKDNIVAQLKOL 439

RESULT 8
OY 096GL3 PRELIMINARY; PRT; 452 AA.
AC 096GL3;
DB 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similar to interferon regulatory factor 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NC NCBI_Taxid=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=BRNIN;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009395; AAH09395.1; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR ProDom; PD002355; IRF.1.
DR PROSITE; PS00601; IRF; UNKNOWN.1.
SQ SEQUENCE 452 AA; 49120 MW; 96B059A028AE6B64 CRC64;

Query Match 13.3%; Score 364.5; DB 4; Length 452;
Best Local Similarity 27.4%; Pred. No. 1.8e-17;
Matches 113; Conservative 43; Mismatches 153; Indels 103; Gaps 16;
OY 10 RVLFGEWMLLGEISSGCGYBGLQWLDARCTCFRVPWGHFARKDLSADARIFKAAVAVGR 69
DB 6 PRIL--PWLVSQDLQGLEGAWVWKSRTFRIPWKGILQDQOEDFGFQAWAERTGA 63
OY 70 WPPSSRGCGPPPEAETARAGKTNFRCAIARSTRFRVWLNDNSGDPADPHKYALSEL 129
DB 64 YVP---GRDPDLPPT-----WKQNFSAIANKREGLRLAEDRSQDPDPKHYEFV--- 110
OY 130 WRBEPGTDQTEAEPAAVPPGPGPFLAHTHAGLQAPGAPLAPAGDGLLLQAVQ 189

```

```

Db      111 -NSGVG---DFSQPDTSPTNGG-----GSTSD-----TQ 136
Qy      190 SCLADHLTASWGDADPPTKAPGEGQGLPLTGACAGGPGLPAGELGMAVETTPSPGPQ 249
Db      137 EDLIDELL-GMMVLAPLP-----DGPFP-----SLAIVAPPCPQ 169
Qy      250 PAALTTGGAAPESPPOAEPLSPSPACTAVOEP-SPGALDVTIMYKGRVLOKVVGH 308
Db      170 PLRSPSLDNPTP-----FPLGSENPPLKRLVPGSEMEFEVYAFRGVFGQTISCP 223
Qy      309 SCTFLYPPDPAPVATPQQVAFPSPA-ELPDQKQLYTEBELRHVAPGLHLEGPQLW 367
Db      224 EQLRLVG-SEVGDRITLPGWPVTLTLPDPMSLTRGCMVSIVRHVLSCLGGGLALMRAGQMLW 282
Qy      368 ARRWGCKVYWEV-----GPPGASPSPTACLLPRNCDPIFDPRVY 410
Db      283 AQRLLGHCHTYWVSEELLPSNGHGPDE-----VPKDKGGEFDLGP 325

```

## RESULT 9

```

Q98TX7 PRELIMINARY; PRT; 445 AA.
ID Q98TX7
AC Q98TX7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Interferon regulatory factor 4.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nenyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
RT Rel-Expressing Fibroblasts."
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF320331; AAK08198.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERREGFCT.
DR PRODOM; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF.1.
SQ SEQUENCE 445 AA; 51120 MW; 7CAE7BDP96780432 CRC64;

```

```

Query Match 13.3%; Score 364; DB 13; Length 445;
Best Local Similarity 26.4%; Pred. No. 1.9e-17;
Matches 130; Conservative 71; Mismatches 201; Indels 90; Gaps 20;
Qy 16 EMLIGESSGCGYGLWLDIARTCPVPMWKFARKDLS-EADARIKRAMAVARGMPSS 74
Db 22 QMLIDQDSGKYPLWVENDKSLFRLPMKAGKQDYRREDALFRAMALFKGKPF--- 77
Qy 75 RGGGPPPEAETAERAGWKTNFRCALSTRFRVMLRDNSG-DPADPHKYVALSRELCSWREG 133
Db 78 REGIDKDPPT-----WKTRLRGALNKSNDPEELVERSQLDISDPKYKVRIVPEGA-KKG 131
Qy 134 PGTDTAEAPAAVPPPGGPPGFLAHTHAGLQAPGLP-----APAGKGLLL 184
Db 132 AKONSMEOQLMNHFFITSP-----YTLSPGVPMVWHERNMBEFAEOPHPDIPY 185-
Qy 185 QAVQOSSCLADHLTASWGDADPPTKAPGEGQGLPLTGACAGGPGLPAGELGMAVETTP 244
Db 186 QGAS-----VPPARQHHWQGGCENGCO-----VTGTTYACAPRESQ 223
Qy 245 SPG-PQPAALTTGGAAPESPPOAEPLSPSPACTAVOEPSPGALDVTIMYKGRVLOK 303

```

```

Db      224 TPGLPIEPSIRSGEALA-----LSDC-----RHICLIYREMLVKEY 260
Qy      304 VGHNP-SCTFLYPPDPAPVATPQQVAFPSPAELPDQKQLYTEBELRHVAPGLHLEJR 362
Db      261 TTSSPFGCRISQ-----QSTEVSLSEQVIFPY--EDNSORKIIEKLHLEGRVILMWA 314
Qy      363 GPOLMARRMGCKVYWEVGGPPGSAS--ESTPACLLPRNCDPIFDPRVFOELVEFRAR 420
Db      315 PDLVYAKRLQOSRIYWD--GPLALGSDRPNK-----LERDQTKLPDTQGFALQLQAF-AH 367
Qy      421 QRRGSPRYTILYGFQDLSAGRPKESLVLVKLEPWLCRVHLEGTREGVSLDSSDL 460
Db      368 HGRPLPRYOVALCFGEFFDPDQ-RQRKLITAHVEPMFARQOLYFAOONSGLHRYGLDPE 426
Qy      481 CUSANSIVDDI 492
Db      427 LMTSPEDYHRSI 438

```

## RESULT 10

```

Q98TX6 PRELIMINARY; PRT; 409 AA.
ID Q98TX6
AC Q98TX6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Interferon regulatory factor 4 deltaB6.
GN IRF-4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAFAS; TISSUE=SPLEEN;
RX MEDLINE=21424622; PubMed=11533227;
RA Hrdlickova R., Nenyba J., Bose H.R. Jr.;
RT "Interferon Regulatory Factor 4 Contributes to Transformation of v-
RT Rel-Expressing Fibroblasts."
RL Mol. Cell. Biol. 21:6369-6386(2001).
DR EMBL; AF320332; AAK08199.1; -.
DR HSSP; P23906; 2IRF.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERREGFCT.
DR PRODOM; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF.1.
SQ SEQUENCE 409 AA; 47489 MW; 0187B3782B62480E CRC64;

```

```

Query Match 13.3%; Score 362; DB 13; Length 409;
Best Local Similarity 25.0%; Pred. No. 2.4e-17;
Matches 123; Conservative 65; Mismatches 178; Indels 126; Gaps 18;
Qy 16 EMLIGESSGCGYGLWLDIARTCPVPMWKFARKDLS-EADARIFRAMAVARGMPSS 74
Db 22 QMLIDQDSGKYGLWVENDKSLFRLPMKAGKQDYRREDALFRAMALFKGKPF--- 77
Qy 75 RGGGPPPEAETAERAGWKTNFRCALSTRFRVMLRDNSG-DPADPHKYVALSRELCSWREG 133
Db 78 REGIDKDPPT-----WKTRLRGALNKSNDPEELVERSQLDISDPKYKVRIVPEGA-KKG 123
Qy 134 PGTDTAEAPAAVPPPGGPPGFLAHTHAGLQAPGLPAPAGDYGLDILLQVQSSCLA 193
Db 124 -----VPGAKKAGKAKONSMEOQLM 143
Qy 194 DHLTASWGDADPPTKAPGEGQGLPLTGACAGGPGLPAGELGMAVETTPSPGPQPAAL 253
Db 144 NH-----PPTITSP-----YTLSPGVPMVW 164
Qy 254 TTGE-----AAAPSPHOAEPLSPS-PSACTAVOEPSPGA-----LDVTIMYKGRVLOK 303

```

```

Db      165 VPERHNRREBAPEDPHDIPYOCASVFAFARGHHMOGPGCENDCRHITCLYYREMLVKEV 224
Qy      304 VVGHP-SCTFLYGPDPBAVATDPOQVAFPSPAELPDQKOLRYTEELLRHVAGLHLEJR 362
Db      225 TTSSPECGRISQ---OSYEVSSLEQVIFYP--EDNSOKRIEKLISLENGVILWMA 278
Qy      363 GPOLMARMGKCKVYWEVGPPGAS--PSTPACLLPRNCDTPIFDPRVFOELVEFRAR 420
Db      279 PDGLYAKRLCOSRIYWD--GPLALCSDRPK---LERDQTKLFTQCFLELQAF-AH 331
Qy      421 QRGSPRTTYLGFQDLSAGRPKESLVLVKLEPMLCRHLEGTQREGVSSLDSDLDL 480
Db      332 HGRPLPRYQVAFCEEFEPDQ--RQRKLITAHVEPMFARQLYYFAQNSGHLRGYDLBE 390
Qy      481 CLASSANSLYDDI 492
Db      391 LMTSPEDYHRSI 402

```

## RESULT 11

```

ID      099419 PRELIMINARY; PRT; 440 AA.
Qy      01-MAY-1997 (TREMBlrel. 03, Created)
DT      01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT      01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE      ICSAT transcription factor (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96239482; PubMed=8657101;
RA      Yamagata T., Nishida U., Tanaka T., Sakai R., Mitani K., Yoshida M.,
RA      Taniguchi T., Iizaki Y., Hirai H.;
RT      "A novel interferon regulatory factor family transcription factor,
RT      ICSAT/PIp/LsIRF, that negatively regulates the activity of interferon-
RT      regulated genes."
RL      Mol. Cell. Biol. 16:1283-1294 (1996).
DR      EMBL: D78261; BA01335.1; -.
DR      HSP: P23906; 21RF.
DR      InterPro: IPR001346; IRF.
DR      Pfam: PF00605; IRF.1.
DR      PRINTS: PR00267; INTERREGCT.
DR      ProDom: PD002355; IRF.1.
DR      SMART: SM00348; IRF.1.
DR      PROSITE: PS00601; IRF.1.
SQ      NON_TER 1
SQ      SEQUENCE 440 AA; 49804 MW; 6A04159FAFBA2701 CRC64;

```

## Query Match

Best Local Similarity 12.6%; Score 344; DB 4; Length 440;

Matches 12; Conservative 60; Mismatches 161; Indels 96; Gaps 21;

```

Qy      16 EWLGEISSGCEGLQWLDEARTCFRVWGHFARKDL--EADARIFKAMAVARGWPSS 74
Db      57 QWLIDQDSGKYPGLWVNEKSIPIPIWGHAKQKQVNRREDALFKAMALFKGR----- 112
Qy      75 RGGPPPEAETAEBAKWTNFRCALRSTRFVMLRDNSG--DPADPHKVYALSRELCSREG 133
Db      113 REGIDKEDPPT-----WKTRLCALNKSNDFEELVERSDLDISDPKYVYIYVEGAKK-- 165
Qy      134 PGTDQTEAEP--AAVPPGSGPPGPEL--ANTHAGIAPGRYLAPAGDKDILLQAVQOS 190
Db      166 -GAKQLTLEDPOKMSHRYMTTTPYSLPAOVHNMMP----- 203
Qy      191 CLADHLLTASWGDVPVTKARGEQEGPLT-----GACAGSPGLPAGELYGMAY 240
Db      204 -----LDRSM-RDYVDORPHRELRYOCSPMTFGRGNHMGSPACENGCQV-TGTFACAP 255
Qy      241 ETPPSFG-PQPAALTTGEAALAPSPHOAEYLSPPSASCTAVOEPFGALDVTIMYKRT 299

```

```

Db      256 PESQAPGVPTSPSIRSAELA-----FSDC-----RLHICLYREIL 292
Qy      300 VLQKVVHP-SCTFLYGPDPBAVATDPOQVAFPSPAELPDQKOLRYTEELLRHVAGLH 358
Db      293 VKELTSSPEGCRISHG--HTYDASNLDQVLFYP--EDNGRKRIEKLISLENGV 346
Qy      359 LELRGPOLMARMGKCKVYWEVGPPGAS--PSTPACLLPRNCDTPIFDPRVFOELVE 416
Db      347 LMAPDGLYAKRLCOSRIYWD--GPLALCNDRPK---LERDQTKLFTQCFLELQAF 400
Qy      417 FRARORRSPRTTYLGFQD 437
Db      401 F-AHGRSLPRFOVTLCEEE 420

```

## RESULT 12

```

ID      090W10 PRELIMINARY; PRT; 416 AA.
Qy      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE      Interferon regulatory factor 10.
GN      IRF10.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPAFAS; TISSUE=SPLEEN;
RA      Nehya J., Hrdlickova R., Bose H.R. Jr.;
RT      "Interferon regulatory factor 10, a novel family member, IRF-10
RT      expression is induced by interferons, concanavalin A, and the Rel/NF-
RT      kappaB oncoprotein v-Rel."
RL      Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF380350; AAK5444.1; -.
DR      InterPro: IPR001346; IRF.
DR      Pfam: PF00605; IRF.1.
DR      ProDom: PD002355; IRF.1.
DR      PROSITE: PS00601; IRF.1.
SQ      SEQUENCE 416 AA; 47646 MW; 669967280FEA967C CRC64;

```

## Query Match

Best Local Similarity 11.8%; Score 321; DB 13; Length 416;

Matches 126; Conservative 65; Mismatches 161; Indels 134; Gaps 27;

```

Qy      16 EWLGEISSGCEGLQWLDEARTCFRVWGHFARKDL--EADARIFKAMAVARGWPSS 74
Db      12 EWLINQDSGRYPGLRWENRERTLPIPKHAKAQDYRQODDALFRAMAVYKGYHEGT 71
Qy      75 RGGPPPEAETAEBAKWTNFRCALRSTRFVMLRDNSG--DPADPHKVYALSRELCSREG 133
Db      72 -----DKADPSYWTTRLCALNKSSTDFQEVERSQLDISPPYKVOI--VC--DG 117
Qy      134 PGTDQTEAEPAAVPPGSGPPGPELANTHAGIAPGRYLAPAGDKDILLQAVQOSCLA 193
Db      118 TRDAEKDEKGRMPTSSKDPQGVAAEESHG-----TAG-----TC-- 154
Qy      194 DHLTASWGDVPVTKARGEQEGPLTGAAGSPGLPAGELYGMAYETTPSPGPAPAL 253
Db      155 -HMPFL-----PLTAPHAE--RGVHYRGI-----FYGM-----SP----- 182
Qy      254 TTGEAALAPSPHOAEYLSPPSASCTAVOEPFGALDVTIMY-----KRTV 300
Db      183 -TRSHLPFRAP-SFLPAEDVNHSDW-----LHIRLYYCDVLVKELTTRTAECRI 231
Qy      301 LQKVVG--HPSCFTLYGPDPBAVATDPOQVAFPSPAELPDQKOLRYTEELLRHVAGLH 358
Db      232 ASRTGPRYGPSCM-----EQIEFPFPAALGGGGMVAVTEVLELDF--H 274

```

QY 359 LEHFGPOLMA-----RRMGCKVYWEVGGP--PGSASBSTPACLLPRNCDFPIEDFRV 409  
 DB 275 LE-RGVLLMVAPEGVEMKROCGGRVYV--NGPLAPHODWENK-----LEREKTYKLLDTQ 327  
 QY 410 FPGELVEFRARQRGRSPRYTYLGFGOD--LSNRPRPEKSLVVKLEPMLCRHVECTOR 467  
 DB 328 FLEBLRRYLS-HGQPAPOYQHLICFGEYFPTSTGRHLQK-LIMAHVEPVARLELPHHAOR 385  
 QY 468 EGVSSL 473  
 DB 386 LGPAL 391

## RESULT 13

Q9QZL7 PRELIMINARY; PRT; 330 AA.

01-MAY-2000 (Tremblrel. 13, Created)  
 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE If66 (Fragment).  
 GN IRF6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/J;  
 RA Sarda A., Mak T.W.;  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF17668; AAF00915.1; -  
 DR HSSP; P23906; 2IRF.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERREGCT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 FT NON\_TER 1  
 FT 330 330  
 SQ SEQUENCE 330 AA; 37355 MW; EB02EC8B751CBD7D CRC64;

Query Match 10.3%; Score 281.5; DB 11; Length 330;  
 est Local Similarity 26.0%; Pred. No. 6.9e-12;  
 atches 103; Conservative 49; Mismatches 161; Indels 83; Gaps 15;

QY 29 GLOWDEARTCFRVPMWGHFAR-KDLSADARIFKAMAVARGMPSSRGSGPPPEAETA 87  
 DB 2 GLIWLHDSKRFQIPWGHATRHSPQEEENTIFKAMAVETGKY---QGVDDPDPAR--- 55  
 QY 88 RAGWKTNFRCAALSTRFVWMLRDNSSG-PADPHKVYLSLELCMBREPGTDQTEAEPAA 146  
 DB 56 ---WKQOLRCALINKSREFNLMYDGTKEVPMNVKIVY---QVC-----D 92  
 QY 147 VPPPGGPPFPFLAHTAGLQAPRLPAPAGDKDLLQAVQSCSLADHLITASWGDVP 206  
 DB 93 IPIQGG-----SVINPSTGSAAPDEKDNVDEDE---EDLEEGSQHHVPI 136  
 QY 207 PTAPAGEGEGRLP---TACAGGRLPAGELYGMAVETTPSPGPQALTTGEAAPE 262  
 DB 137 QDTFFPFININGSPPAPASVGNCSVGNCSF-----ESVWPKTEPLEMEVPOAPI 184  
 QY 263 SPHOAEYLLSPSPACAVQSPGALDVTIMYKGRVLOKV-VGHP-SCFTFLYGPDP- 319  
 DB 185 QPFTVSSELMTSSLPMT-----DLDIKYQKRGKYGQGMVTSNPGCRLEFGDIGPM 236  
 QY 320 -----AVRATDPQOAVFPSPAPLPDQOLRYTEBELRHVAGLLELRGQPLMARMGKC 374  
 DB 237 PDQELRGVPSLEQVKFPFGPEHINKEOKLFTSLDLVMDRGILLLEVSGHAIVAIRLCQC 296  
 QY 375 KYVWEVGGPGSASPSPTACLLPRNCCTPIEDFRVF 410

DB 297 KYWSGPCASLAPN-----LIERQKKVLFLETF 328

## RESULT 14

Q924T6 PRELIMINARY; PRT; 215 AA.

01-DEC-2001 (Tremblrel. 19, Created)  
 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 GN ISGF3G.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=21243724; PubMed=11345588;  
 RA Yawata M., Murata S., Tanaka K., Ichigatsubo Y., Kasahara M.;  
 RT "Nucleotide sequence analysis of the ~35-kb segment containing  
 RT interferon-gamma-inducible mouse proteasome activator genes";  
 RL Immunogenetics 53:119-129(2001).  
 DR EMBL; AB053120; BAB47407.1; -  
 DR MGI; MGI:107587; Isgf3g.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRODOM; PD002355; IRF.1.  
 DR PROSITE; PS00601; IRF; UNKNOWN\_1.  
 SQ SEQUENCE 215 AA; 24028 MW; 5C2004D507F0C64D CRC64;

Query Match 7.1%; Score 194; DB 11; Length 215;  
 Best Local Similarity 35.5%; Pred. No. 4.7e-06;  
 Matches 39; Conservative 21; Mismatches 38; Indels 12; Gaps 3;

QY 17 WLGEISSGCGYEGLOWDEARTCFRVPMWGHFARKDUSE-ADARIFKAMAVARGMPSSR 75  
 DB 15 WIVEQVESGHFPVCDDAKTIFRIPMWGHAGQDFREODDAIFAMALFKK----- 68  
 QY 76 GGGPPPEAETAEKAGKTNFRCAALSTRFVWMLRDNSSG-DPADPHKVYAL 124  
 DB 69 ---HKDGDIGHPAVWKTRLCALINKSSSEFEVPERGRMDVAPRYVYRI 114  
 RESULT 15  
 Q9KRK9 PRELIMINARY; PRT; 1334 AA.  
 AC Q9KRK9;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Putative multi-domain regulatory protein.  
 GN SC02259 OR SC075A.05C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA MEDLINE=97000351; PubMed=8843436;  
 RX Redenbach M., Kiese H.M., Denapate D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL133220; CAB61705.1; -  
 DR InterPro; IPR005158; BAD  
 DR InterPro; IPR000767; Disease\_resist.  
 DR InterPro; IPR001867; Trans\_reg\_C.  
 DR Pfam; PF03704; BAD; 1.  
 DR PRINTS; PR00364; DISEASERISIT.  
 DR PRINTS; PR01574; TUBBYPROTEIN.  
 DR Prodom; PD000329; Trans\_reg\_C; 1.  
 SQ SEQUENCE 1334 AA; 138787 MM; 78DC746883E8778C CRC64;

Query Match 6.7%; Score 183; DB 16; Length 1334;  
 Best Local Similarity 28.1%; Pred. No. 0.00021;  
 Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;

QY 18 LIGEISGCTEGLOMDBEATCRVPMKHFARKDLSADARIFKAMAVARGMPSSRG 77  
 DB 146 VLADLPDRTAEARW--ETHFEALRARTALDLGOEHSLELTALCDG----- 195  
 QY 78 GPPPEAETAE-----AGWKTNRFCALRSTRFRFVMDRNSG--DPA-----D 117  
 DB 196 -PLDEPLQALRLRALRDSGRTAEALAYEAVRR--LIADRIGTDPGPELRTLHAEILSPS 252  
 QY 118 PHKVVALSRRLCWREGP-----GTDQTEAEPAAVPPQGP-PGPFLAHTHAGL 166  
 DB 253 PTPTPGRSRTPGWTSGCPASGAGAAGTIDVAGAGASGDPDASGPASGPAVAPGSGG 312  
 QY 167 QAGGELPAPAGDGDLLQAVQOSCLADHLLTASWGADPVPTKAPGEGEGLPLTGACAG 226  
 DB 313 PARGMWPAPGTAPGSSSTAPPHDTASAD--TA--PAPGPTAPGTA--PAGTAA 362  
 QY 227 GPGI--PA-GELY--GWA--VETTPSPGPQPAALTTGEA-----AAPSPHOAEPYL 271  
 DB 363 APGTAGAPGTAAGTAPGTAAGTAPGTAAGTAPGTAAGTAPGTAAGTAPGTAAGTAPGTA 422  
 QY 272 SPSPSACTAV-----QEPSGALDVTIMYKGRIVLQKVGHPSCTFLGPPDPAVRAT 324  
 DB 423 APAGSTPAGTVPAPGTAPAGTAPGPA--DGR--RPVTPGASG--GFG--AAAT 468  
 QY 325 DPQOVA-----FPSPA 335  
 DB 469 PPEAAAAASAGSAPSPA 485

Search completed: June 18, 2003, 12:47:49  
 Job time : 39 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:45:13 ; Search time 22 Seconds  
(without alignments)  
2197.982 Million cell updates/sec

Title: US-09-647-965-9  
Perfect score: 2731  
Sequence: 1 MALAPERARVLFGEMLLG.....SANSLYDDIECFIMLEQPA 503

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	780	28.6	491	2	S56753
2	482.5	17.7	504	2	G02474
3	440	16.1	459	2	UC6520
4	376	13.8	450	2	S57837
5	333.5	12.2	424	2	A35861
6	328.5	12.0	393	2	A45017
7	306	11.2	425	2	A45064
8	259.5	9.5	399	2	UC4592
9	224	8.2	203	2	S57836
10	183	6.7	1334	2	T50568
11	176	6.4	325	2	I52998
12	176	6.4	1446	2	A38587
13	167	6.1	349	2	A53340
14	166	6.1	325	2	B31595
15	166	6.1	329	2	A31595
16	165.5	6.1	328	2	A36330
17	160.5	5.9	1400	2	T31555
18	160	5.9	1747	1	A45974
19	159.5	5.8	416	1	SKXIAG
20	159	5.8	576	2	T36729
21	158.5	5.8	839	2	T04859
22	157.5	5.8	705	2	A35363
23	156.5	5.7	1464	2	S59856
24	156	5.7	381	2	T27806
25	156	5.7	1857	2	S31212
26	156	5.7	1888	2	S78476
27	156	5.7	3149	1	Q0858
28	155	5.7	801	2	T29018
29	154	5.6	660	1	Q0853

30	154	5.6	3530	2	A59266	unconventional myo
31	150.5	5.5	691	2	A25704	synapsin I - rat
32	150	5.5	704	2	A30411	synapsin Ia - rat
33	149.5	5.5	1952	2	T48814	hypothetical prote
34	148.5	5.4	560	1	WPHUM	multimeric inhibiti
35	148	5.4	296	2	H70580	hypothetical prote
36	148	5.4	1464	1	CGHUI5	collagen alpha 1(I
37	148	5.4	1466	1	CGHUI7	collagen alpha 1(I
38	147	5.4	301	2	T21314	hypothetical prote
39	147	5.4	633	2	B40983	collagen alpha 1(X
40	146.5	5.4	1042	1	CGHIS	collagen alpha 1(I
41	145	5.3	296	2	A31219	collagen 1 - Caeno
42	145	5.3	1207	2	T00378	KIAA0641 protein -
43	145	5.3	2715	2	T13049	eyelid - fruit fly
44	144.5	5.3	1315	2	A56101	collagen alpha 1(X
45	144.5	5.3	1546	1	CGHU2E	collagen alpha 2(X

## ALIGNMENTS

## RESULT 1

S56753  
Interferon regulatory factor 3 - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: S56753  
R:Grant, C.E., Vasa, M.Z., Deeley, R.G.  
Nucleic Acids Res. 23, 2137-2146, 1995  
A:Title: CIRP-3, a new member of the Interferon regulatory factor (IRF) family that is  
A:Reference number: S56753; MUID:95334365; PMID:7541908  
A:Accession: S56753  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-491 <GRA>  
A:Cross-references: EMBL:U20338; NID:G790580; PIDN:AA86995.1; PID:G790581  
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 28.6% Score 780; DB 2; Length 491;  
Best Local Similarity 38.5%; Pred. No. 3.5e-40;

Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY	2	ALAPERARVLFGEMLLG	ISGCYEG	LOWLDEARTCFRVPWK	FARKDISEADARIFK	61
DB	3	ALDSEGDQKURFGPMLNA	VSSGLYRGLCMIDPDRRI	FRIPWKNARKDVTSS	SVETLIFK	62
QY	62	AMAVARGRWPSSRGSGP	PEAEIARAGWKTFRCAL	RSTRFVMLDNGSDPAD	PHKV	121
DB	63	AMAKASGRV	-----EGNADPAKMKTN	FRCALRSTHFMIL	EDRSKCNDPPHKV	111
QY	122	VALSRELCMRECPGT	DTDEARPAVPPQ	---GSP	-----PGPFLAHNTAGQ	167
DB	112	YAVA	-----SGVPNDRSGG	VAGALQOQPOLLN	HHDLAENTPT	152
QY	168	APGFLPAPAGDKGL	LLQAVQCSCLADHLL	TASWAGDPVPTKAPGEG	---QEGPLTG	222
DB	153	DSTEGVAAALQVLDL	DSVLOCHCNISAL	-----GQPTIMATGAL	PBDALLPQ	207
QY	223	ACAGPGPLPAGELYG	AVETTPSPPOP	-----AALTGEAADES	PHQAP	269
DB	208	DCCLTGP	-----QFQDWRLQLEP	PLLGQPLTGCGCG	DGAGALPVSCEALPAPSPAE	262
QY	270	YL	-----SPSPACTAVQ	SPSGALDVTIMYGR	TVLQVYGHPSCTFLYGPDP	VRATD 325
DB	263	LLFGSANPAPPPAD	IGLPLDITTYRGKMYQ	QOVDSRCVLAVYQ	PDPAV--AB	320
QY	326	POQVAFSPAPAE	LDQKQRYTEELLRHVAP	GLHLELRGPOLMAR	BMGCKVYEW	380
DB	321	QRLVLPSPAPAS	PDQRQRYTEDLLF	--VAGRLRQGRAG	QLLATLTKCKCKVFWALSQ	378
QY	381	GSPGSASBPST	PACLLPNCDDTPI	PDFRVFQELVEFPAR	QRSGSPRYTIYGFQ	QDLSA 440
DB	379	GSEP	-----PLNLHRDQ	ETITIPDFRVFCTEL	RDFDRSRERSRSPDFTIFL	CFQCCSS 431





QY 76 GGGPPPEAETAEERAGMKTNFPCALSTRFRFVWLRLDNGS-PPADPHKYVALSRELQWREGP 134  
 Db 70 -----KSGDTGGRVAVWTKRLCALINKSSEFEKEVERGRMDVAEYKYQLLPPIVSGOP 124  
 QY 135 GTDQTEAEAPAAVPPPOGGPPPLAHTAGLQAPGLPAPADKGLLLQAVQSCGLAD 194  
 Db 125 GTQKVPSPK-----ROHSVS-----SRKEED-----AMQNTLSP 156  
 QY 195 HLLTASGADPVPTKARAGEQGLPLTGACAGSGPLPAGELYGNAVETTPSGOPALYT 254  
 Db 157 SVLQDSLNNEE-----EGASG-----GAHSDIG-----SSSSSSPPEQEVTD 196  
 QY 255 TGEAAAPESPHQAE-----PYLSPSPSACTAVQEPSPG-ALDVTIMYGRVTLQKVGHP 309  
 Db 197 T-----EAPFGQGRRLLEFLP-----PEPYSLTLITFYNGRVVGEKQVOSLD 240  
 QY 310 CTPLYPGPPDPAVATPDQVAFSPSPALPDQKQRLYTEELLRHVAPGLHLRGPOLMAR 369  
 Db 241 CRUV-----AEPSSGSESSMEQVLFKPGPLEP-----TQRLLSQLERGILVANSNPGLFVQ 291  
 QY 370 RMGCKKYWEVGGPGSASPTACILPRNCDTPIFPFRVFPQVLVFRARQRRG-SPRY 428  
 Db 292 RLCPPIPSWAPQAPPGPGPH-----LLPNSCEVELFTAYFCRDLV--RYFQGLGPPPKF 345  
 QY 429 TIYLGFGQDLSAGRPKESLVLVLEPMLCRVHLEGT--OREGVSL 473  
 Db 346 QVTLNPFESHSGSHTPQNLITYMEOAFARYLLEQTPQOQALSL 392

## RESULT 7

A45064  
 Interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human  
 C/Species: Homo sapiens (hmn)  
 C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Sep-1998  
 C/Accession: A45064  
 R/Weisz, A.; Marx, P.; Sharif, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levy, B.Z.  
 J. Biol. Chem. 267, 25589-25596, 1992  
 A/Title: Human interferon consensus sequence binding protein is a negative regulator of  
 A/Reference number: A45064; MUID:33094284; PMID:1460054  
 A/Accession: A45064  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-425 <MEI>  
 A/Experimental source: lung and blood  
 A/Note: sequence extracted from NCBI backbone (NCBI:P.120312)  
 C/Superfamily: lymphoid-specific interferon regulatory factor  
 C/Keywords: DNA binding; transcription regulation

Query Match 11.2%; Score 306; DB 2; Length 425;  
 Best Local Similarity 23.4%; Pred. No. 1.6e-11;

Matches 121; Conservative 67; Mismatches 174; Indels 156; Gaps 21;

QY 16 EMLLGESGCGYGLQMLDPAITCFRVPWKHPARKDIS-EADARIFAMAVARGMPSSR 74  
 Db 12 QWLEIQDIDSMYPGLIMENEKSMFRIIPWKGAGKQDQNVDSIFKAMAVFKGK----- 67  
 QY 75 RGGGPPPEAETAEERAGMKTNFPCALSTRFRFVWLRLDNGS-PPADPHKYVALSREL 128  
 Db 68 -----KSGDAEAPWTKRLCALINKSPFEETDTSQDLISBPYVYVIVVEEDQCK 121  
 QY 129 -----CMREGGTDQTEAEAPAAVPPPOGGPPPLAHTAGLQAPGLPAPADKGLD 181  
 Db 122 LGVATAGCVNEVMEGRSEIDELIKERS-----VDDYGMKTRSSSPEDA----- 168  
 QY 182 LLIQAVQSCGLADHLLTASGADPVPTKARAGEQGLPLTGACAGSGPLPAGELYGNAV 241  
 Db 169 -----CRSQ--LLPDM-----WAHE 181  
 QY 242 TTPSPGPAPALTTGEAAAPESPHQAEPYLSPSPSACTAVQEPSPGALDVTIMYKGRTVL 301  
 Db 182 --PSTGR--LVYGTITY--DAHSAF-----SOMVISFYGGKLVG 217

QY 302 QKVVGHP-SC-----TFLYGPPDPAVATPDQVAFSPSPALPDQKQRLYTEELL 349  
 Db 218 QATTCPEGGRSLSLSGPLGPTLYGP-----EDELIVRFPAPATISSERQVTRKL 270  
 QY 350 LRHVAPG-LHLERGPOLWARRMGCKKYWEVGGPGSASPTACILPRNCDTPIPDR 408  
 Db 271 FGHLERGVLIHSSRQG--VFAVKRLQGRVFCSGNAVVCCKGRPNK-----LRDEVVQVFDTS 324  
 QY 409 VFQELVERPARRRSPRTIYLGFGQDLSAGRPKESLVLVLEPMLCRVHLEGT 468  
 Db 325 QFFRELQGFYNSQGR-LPGRVVLCPGEEPPDAPLRSKLILVQIE-----QLYROLAEB 379  
 QY 469 GVSSLDSDL-----DLCLSSANSLY 489  
 Db 380 AGKSCAGSYMQAPPEPPPDQVRMPEDICASHQRSEFF 417

## RESULT 8

JC4592  
 transcription factor ISGF3 gamma chain - mouse  
 N/Alternate names: interferon-stimulated gene factor 3 gamma chain  
 C/Species: Mus musculus (house mouse)  
 C/Date: 10-Apr-1996 #sequence\_revision 24-Oct-1998 #text\_change 21-Jul-2000  
 C/Accession: JC4592; S71599  
 R/Suhara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.  
 J. Biochem. 119, 231-234, 1996  
 A/Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma/p48)  
 A/Reference number: JC4592; MUID:97037063; PMID:8882710  
 A/Accession: JC4592  
 A/Molecule type: mRNA  
 A/Residues: 1-399 <SUB>  
 A/Cross-references: EMBL:U51992; NID:g1263309; PIDN:AAC52494.1; PID:g1263310  
 A/Experimental source: L929 cells  
 R/Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, M.  
 FEBS Lett. 358, 225-229, 1995  
 A/Title: Possible involvement of the transcription factor ISGF3gamma in virus-induced ex  
 A/Reference number: S71599; MUID:95145714; PMID:7843405  
 A/Accession: S71599  
 A/Status: not compared with conceptual translation  
 A/Molecule type: mRNA  
 A/Residues: 1-87 <KAW>  
 A/Gene: Isf3g  
 A/Genetics: 1-87 <KAW>  
 A/Map position: 14  
 C/Function:  
 A/Description: responsible for specific interaction with the promoter element, interfere  
 ed gene factor 3, the primary regulator of type I interferon responses; involved in type  
 A/Note: Induced by interferon-alpha and interferon-beta  
 C/Superfamily: lymphoid-specific interferon regulatory factor  
 C/Keywords: DNA binding; signal transduction; transcription factor  
 F/10-117/Domain: DNA binding #status predicted <DNB>

Query Match 9.5%; Score 259.5; DB 2; Length 399;  
 Best Local Similarity 22.5%; Pred. No. 9.6e-09;

Matches 105; Conservative 64; Mismatches 192; Indels 105; Gaps 14;

QY 17 EMLLGESGCGYGLQMLDPAITCFRVPWKHPARKDIS-EADARIFAMAVARGMPSSR 75  
 Db 15 WYEOVESGHFPVCCDDDAKTMFRIIPWKGAGKQDQNVDSIFKAMAVFKGK----- 68  
 QY 76 GGGPPPEAETAEERAGMKTNFPCALSTRFRFVWLRLDNGS-PPADPHKYV-----ALSREL 128  
 Db 69 -----HKGDIIGHPAVWTKRLCALINKSSEFEVEPERGRMDVAEYKYVILPAGTLPNCP 124  
 QY 129 CMREGGTDQTEAEAPAAVPPPOGGPPPLAHTAGLQAPGLPAPADKGLLLQAVQ 188  
 Db 125 RNOKSPCKRSISCVSFERENMENENGRITGVVNHSSGSNIGG-----GNGS----- 171  
 QY 189 QSGCLADHLLTASGADPVPTKARAGEQGLPLTGACAGSGPLPAGELYGNAVETTPSPGP 248  
 Db 172 -----NNSDSNCSNSELDEG-----AG----- 189  
 QY 249 OPALTTGEAAAPESPHQAEPYLSPSPSACTAVQEPSPGALDVTIMYKGRTVLQKVVGHP 308

Db 190 -----TTTATREDVLEHQLPLN-----SDYSLLTFTYGGVAVKQTVHSL 233

Qy 309 SCTFLYGPDPRAVRAIDPOQVAFPPAPBLBDQKQALYTELLAHVAPGLHLELRGQVLA 368

Db 234 DCLVLAERSD-----SESSMEQVEFPKP-----DPLEPTQHLNLQLRGVLAASNSGLV 284

Qy 369 RRMGCKVYWEVGGPPGASAPSTPACLLPNCCTPIFDFFVFPQELVEFRARQRG-SPR 427

Db 285 QRLCEPFIQWNAEAPRGPBP-----LIPSNKCEVEFKTTFRCRDLAQY--FGCGGPBP 338

Qy 428 YTIYVGFQDLSAGRPKESLVVLKLEPMLCRVHLEGTQREGVSSL 473

Db 339 FOATLHFWEESPGSSHQENLITVQMEQAFARHLEKIPPEEKAL 384

## RESULT 9

836  
Lymphoid-specific interferon regulatory factor - mouse (fragments)  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
C/Accession: S57836  
R/Matsuura, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Klefer, F.; Kawakami, Nucleic Acids Res. 23, 2127-2136, 1995  
A>Title: Molecular cloning of LSIF, a lymphoid-specific member of the interferon regulatory factor family  
A/Reference number: S57836; MUID:95334364; PMID:7541907  
A/Accession: S57836  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-203 <MAT>  
A:Cross-references: EMBL:U11692  
C:Superfamily: lymphoid-specific interferon regulatory factor

Query Match 8.2%; Score 224; DB 2; Length 203;  
Best Local Similarity 34.8%; Pred. No. 6.7e-07;  
Matches 57; Conservative 22; Mismatches 49; Indels 36; Gaps 6;

Qy 16 EMLTGISGCGEGLQWLDEARTCFRVPWKHFARKDL--BADARIFKAMAVARGMPSS 74

Db 4 QWILDQIDSKYGLWNEKEKVFRIPMKAGKQDNRREDALFRVAMLPFGKPF----- 59

Qy 75 RCGGPPPEAETARAGKTNFRCALSTRFRVLRDNG--DPADPHKVALSRELQWREG 133

Db 60 REGIDKRPDPT-----WKTRLCALNKSNDFEELVRSQDIDPYKVRNI----- 105

134 PGTDQTEAERPAVPPPGGPPFPFLAHTAGL---QAPPLPA 174

Db 106 -----VPEQQRNIEKLSHLERGLVLMWAPDGLYA 136

## RESULT 10

T50568  
Probable multi-domain regulatory protein [imported] - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C/Accession: T50568  
R/Redenbach, M.; Kiese, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopx, M.; Microbiol. 21, 77-96, 1996  
A>Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S. coelicolor chromosome  
A/Reference number: T20556; MUID:97000351; PMID:8843436  
A/Accession: T50568  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1334 <RED>  
A:Cross-references: EMBL:ALJ33220; PIDN:CAB61705.1  
A:Experimental source: strain A3(2)  
C:Genetics:  
A>Note: SCC75A.05c

Query Match 6.7%; Score 183; DB 2; Length 1334;  
Best Local Similarity 28.1%; Pred. No. 0.0015;  
Matches 106; Conservative 25; Mismatches 150; Indels 96; Gaps 22;

Qy 18 LUGELISSGCGEGLQWLDEARTCFRVPWKHFARKDLSEADARIFKAMAVARGMPSSRG 77

Db 146 VLADLPDRTAEARW--ETRHFEALRARTALDQGAHSH;PELTALCDG----- 195

Qy 78 GPPPEAETAR-----AGWKTNFRCALSTRFRVLRDNG--DPA-----D 117

Db 196 -PLDEPLQALRLRALRDSGRTAEALAAVEAVRR--LLDRLCTDQDPBELRTLHAEILSPS 252

Qy 118 PKVVALSRELQWREGP-----GNDQTEAERPAVPPPGGPP--PGFLAHTAGL 166

Db 253 PTPTRGRSTPPTWTGTPGPASAGAASTDVASGAAAGPPPASGAPVAVAGSGGG 312

Qy 167 QAPGVPAPAGDKDULLQAVQSCILADHLITASWAGDPVPTKABEGQEGPLTGACAG 226

Db 313 PAPGWWPAPGTAAGSTAPPHDTASAD--TA-----PAPGPTAPGTA---PAGTAAP 362

Qy 227 GPGL--PA-GELY--GWA--VETTSPPGPPALTTGEA-----AAPSPHOAEPL 271

Db 363 AGTAGPAPAGTAYAGTAPVAGCTTAPGTAAPGTAAPGARDTSYAGTAPVAGTTAPAGT 422

Qy 272 SPSPACTAV-----QEPSPALDVTIMYGRITVLOKVGHPACTFLYGPDPAVRAT 324

Db 423 AAPGSTAPAGTVPAPGTAAPAGPQPA-----DGR---RVTGPASGT---GFG---ANT 468

Qy 325 DPOQVA-----FPSPA 335

Db 469 PPEAAAAAGSAPSPA 485

## RESULT 11

152998  
Interferon regulatory factor 1 - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C/Accession: I52998  
R/Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deisseroth, A.B. DNA Cell Biol. 11, 605-611, 1992  
A>Title: Human interferon regulatory factor 1: intron/exon organization.  
A/Reference number: I52998; MUID:93000481; PMID:1382447  
A/Accession: I52998  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-325 <RES>  
A:Cross-references: GB:I05072; NID:G184648; PIDN:AAA36043.1; PID:G184649  
C:Genetics:  
A:Gene: GDB:IRF1  
A:Cross-references: GDB:127269; OMIM:147575  
A:Map position: 5q31.1-5q31.1  
A:Intons: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1

Query Match 6.4%; Score 176; DB 2; Length 325;  
Best Local Similarity 23.6%; Pred. No. 0.00089;  
Matches 69; Conservative 35; Mismatches 108; Indels 80; Gaps 12;

Qy 11 RVLFGMLTGISGCGEGLQWLDEARTCFRVPWKHFARK--DLSEADARIFKAMAVARG 68

Db 5 RRMRRWLMQINSNDIPGLIMINKEMEFQIPMKGAAGHGDINK-DACLRSAWIHTG 63

Qy 69 RMPSSRGCGPPPEAETARAGKTNFRCALSTRFRVLRD--NSGPDADPHKYVALS 125

Db 64 RY-----KAEKEFPDPT-----WKNFRCAMNSLPDIEVKQSRNKGSSA--VRYYRM 112

Qy 126 RELC---WREGTQDTEAERPAVPPPGGPPPP-----LAHTAGLQAPPLPA 174

Db 113 PPLTKNQRERKSKSRDASKAKRSCDSSPDFTSDGLSSSTLPDHSSTVPQ----- 168

Qy 175 PAGDKGDLLOAVQSCILADHLITASWAGDPVPTKABEGQEGPLTGACAGGPGP--- 231

Db 169 -----YMQDLVEQALT-----PALSPCAVSTLPDMH 196

Qy 232 -----AGELYGAVAVETTPSPGPPALTTGEAAPSPHOAEPLVSPS 274

Db 197 IPVEVVPDSTSDLYNFQV-----SPMSTSEATTDDEBGLPEDIMKLEOS 244

## RESULT 12

A38587  
 A:Species: cornea-specific - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 14-Feb-1992 #sequence\_revision 15-Aug-1997 #text\_change 20-Sep-1999  
 C/Accession: S16501; A38587  
 R:Merchant, J.K.; Linsemayer, T.F.; Gordon, M.K.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 1560-1564, 1991  
 A>Title: cDNA analysis predicts a cornea-specific collagen.  
 A:Reference number: A38587; MUID:91142213; PMID:1705041  
 A/Accession: S16501  
 A:Molecule type: mRNA  
 A:Residues: 1-1146 <MAR>  
 A/Cross-references: EMBL:M60172; NID:9211609; PIDN:AAA48703.1; PID:9211610  
 A/Accession: A38587  
 A:Molecule type: mRNA  
 A:Residues: 1-174,'X',176-223,'X',235-344,'X',346-408,'X',410-499,'X',501-876,'X',878-114  
 A/Cross-references: GB:M60172  
 C:Superfamily: unassigned collagens  
 C:Keywords: cornea

Query Match 6.4%; Score 176; DB 2; Length 1146;  
 Best Local Similarity 23.8%; Pred. No. 0.0033;  
 Matches 109; Conservative 31; Mismatches 132; Indels 186; Gaps 26;

QY 66 ARGWPPSSRG-----GPPPEAETAEBAKMTNFRCLSTRFRVWLKRDNSGPPAD 117  
 DB 348 ARGPPSGSGTGERGLTGPQGPPLPGNPGRPKAK-----GPPGA 387  
 QY 118 PHKYALSRLELCWEGPGTQTEAEAPAAVPPGPGPPPLATHAGLOAP-GP-LP- 173  
 DB 388 PKGVIS-----AECSPTIALRGP-GRPGPGPGPPGPPGPPADLPG 431  
 QY 174 --APAGKX--DILLQAV--QOSCLADHLITASWG--ADVPYKADGEGGLPLTGAC 224  
 DB 432 QQGRGKSGSAVEVIEITIKTEVSSLASOMLSDQGRAGPPGPPGSGSVQGLP--GP 488  
 QY 225 AGPGGLPAGELYGAVETTS--PGPPALLTGE--AAPESP-----HQAEPYLS 272  
 DB 489 RGPGLG-----PSGPPGPPSGSVSTSETVSGPPGPPGPPGKDDGE--- 533  
 QY 273 PPSACTAVQEPSPGALDVTIMYKRTVLQKVGHPSCTFLYGPDPVAVATDP----- 326  
 DB 534 PGRRGFTG--EPGPPGLPGRSSHGCTYTMQGPPEP-----GPPGKGAAGVAGARIP 585  
 QY 327 -----QGV-----APSPAPLPPQKQLRYTEELI-----RHVAPGLHLELRGP 364  
 DB 586 GTSRSGSRQIQGPPGPPGPPGPPGCGSSQDEIQQYVADYLDKSDNVHLYLTG----- 637  
 QY 365 QLMARNGKCKVWVEGPPGSGASPSPPACLPNCCPTPIPDF----- 407  
 DB 638 -----VGPPG--PPGPPGIL--TTAAGKNDFEALATRWMSVTSSDH 678  
 QY 408 -----RVFQELVEFRARQ-----RGSP 426  
 DB 679 YQSFASSVSTSVLYQELLMNLQGEELRQYLVGRGPP 716

## RESULT 13

A53340  
 A:Species: Homo sapiens (man)  
 N:Alternate names: transcription repressor IRF2  
 C:Species: Homo sapiens (man)  
 C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 05-Nov-1999  
 C/Accession: A53340; S06894; A32828  
 R:Chen, Y.; Deisseroth, A.B.  
 J. Biol. Chem. 269, 5279-5287, 1994  
 A>Title: Human interferon regulatory factor 2 gene. Intron-exon organization and function.  
 A:Reference number: A53340; MUID:94148994; PMID:8106512  
 A/Accession: A53340  
 A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-349 <CHA>

A/Cross-references: GB:L24442

R:Ich, S.; Harada, H.; Fujita, T.; Mimura, T.; Taniguchi, T.

Nucleic Acids Res. 17, 8372, 1989

A>Title: Sequence of a cDNA coding for human IRF-2.

A:Reference number: S06894; MUID:90045964; PMID:2813069

A/Accession: S06894

A:Molecule type: mRNA

A:Residues: 1-57,'R',59-349 <ITO>

A/Cross-references: EMBL:X15949; NID:933966; PIDN:CAA34073.1; PID:933967

R:Harada, H.; Fujita, T.; Miyamoto, M.; Kimura, Y.; Maruyama, M.; Furuta, A.; Miyata, T.

Cell 58, 729-739, 1989

A>Title: Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to

A:Reference number: A32828; MUID:89354547; PMID:2475256

A/Accession: A32828

A:Molecule type: mRNA

A:Residues: 1-68,'T',70-96,'R',98-130,'ER',133-152,'GF',155-163,'A',165-188,'D',190-210,

8,'T',310-313,'PAPV',318-319,'TP',322-349 <HAR>

A/Cross-references: GB:J03168; NID:9198455; PIDN:AAA9333.1; PID:9293676

A:Experimental source: clones 2 and 5

C:Genetics:

A:Gene: GDB:IRF2

A/Cross-references: GDB:127270; OMIM:147576

A:Map position: 4q35.1-4q35.1

C:Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 167; DB 2; Length 349;  
 Best Local Similarity 35.0%; Pred. No. 0.0033;  
 Matches 36; Conservative 18; Mismatches 39; Indels 10; Gaps 3;

QY 11 RVLFQWMLIGELSSGCEGLQWLDERTCFRVPMKHFARKDLS-EDARIFKAAVARGP 69  
 DB 5 RMMRMWLEEQINSNTIPGLKMLNKEKTFQIPWMAAHHGMVDEADALFRWALHTGK 64  
 QY 70 WPPSSRGGPPPEAEFAERAGKMTNFRCLSTRFRVWLKRDNS 112  
 DB 65 HQP-----GVDKPDPKT-----WRANFRCAMNSLPDIEEVDKDS 98

## RESULT 14

B31595  
 A:Species: Homo sapiens (man)  
 C:Species: Homo sapiens (man)  
 C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 16-Feb-1997  
 C/Accession: B31595; S04075  
 R:Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.  
 Cell 54, 903-913, 1988  
 A>Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specific  
 A:Reference number: A90903; MUID:88311092; PMID:3409321  
 A/Accession: B31595  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-325 <M12>  
 R:Maruyama, M.; Fujita, T.; Taniguchi, T.  
 Nucleic Acids Res. 17, 3292, 1989  
 A>Title: Sequence of a cDNA coding for human IRF-1.  
 A:Reference number: S04075; MUID:89263736; PMID:2726461  
 A/Accession: S04075  
 A:Molecule type: mRNA  
 A:Residues: 1-325 <MAR>  
 A/Cross-references: EMBL:X14454  
 C:Keywords: DNA binding; transcription regulation

Query Match 6.1%; Score 166; DB 2; Length 325;  
 Best Local Similarity 23.4%; Pred. No. 0.0036;  
 Matches 67; Conservative 33; Mismatches 106; Indels 80; Gaps 12;

QY 17 WLGEISSGCEGLQWLDERTCFRVPMKHFARK--DLSEADARIFKAAVARGPSS 74  
 DB 11 WLMQINSNOIPGLINKEMILEIPWMAAHHGMVDEADALFRWALHTGK 65  
 QY 75 RGGPPPEAETAEBAKMTNFRCLSTRFRVWLKRD--NSGDPADPHKYALSRLELC-- 129

```
Db      66 KAGEKEPDPTK-----WKANFRCAMNSLPDIEVKQSRNKGSSA--VRYYRMLPPLTKN 118
QY      130 -WREGPGTDQTEAEAPAAVPPQGGPPGP-----LAHTHAGLQAPGFLPAPAGDKG 180
Db      119 QRKERKSKSRDASKKAKKSCGDSPTDSSSTLPDDHSSYTVPG-----168
QY      181 DLLQAVQSCGLADHLITASWGADPVPTKAFGEQEGPLTGACAGGPGLP-----231
Db      169 -----YMQLVEVQALT-----PALSPCAVSTLPDMHIPEVV 202
QY      232 ---AGELYGWAVENTPSPGPQPAALTGEAAAPSPHOAEPIYSPS 274
Db      203 PDSTSDLYNFOV-----SPMPSISEATDEDEKGLPEDIMKLEQS 244
```

## RESULT 15

```
Interferon regulatory factor 1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 05-Nov-1999
C/Accession: A31595
R/Miyamoto, M.; Fujita, T.; Kimura, Y.; Maruyama, M.; Harada, H.; Sudo, Y.; Miyata, T.;
Cell 54, 903-913, 1988
A/Title: Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifies
A/Reference number: A90903; MUID:88311092; PMID:3409321
A/Accession: A31595
A/Molecule type: mRNA
A/Residues: 1329 <MTY>
A/Cross-references: GB:M21065; NID:g198458; PIDN:AAA39334.1; PID:g293677; GB:J03160; GB:
C/Keywords: DNA binding; transcription regulation
```

## Query Match

6.1%; Score 166; DB 2; Length 329;

Best Local Similarity 21.6%; Pred. No. 0.0036;

Matches 65; Conservative 40; Mismatches 88; Indels 108; Gaps 12;

```
QY      11 RVLFGEWLLGEISSGCEYGLQWLDEARTCFRVPWKHPARK--DLSEADARIFKAAVARG 68
Db      5 RMRRPWLWEMQINSNOIPLIWKEMIFQIPKHAKHGWDINK--DACLFPSMAIHTG 63
QY      69 RWPSSSRGGPPPEAETAEERAGWKTNFRCAIRSTRRFVMLRD--NSGDPADPHKYAL- 124
Db      64 RY-----KAGEKEPDPTK-----WKANFRCAMNSLPDIEVKQSRNKGSSA--VRYYRML 112
QY      125 -----SRELCPWREGPGT-----136
Db      113 PPIURNQRRKRSKSSRDTSKTKRKLCDVSPDTPSDGLSSSTLPDDHSSYTTQGYLQ 172
QY      137 -----DQTEAEAPAAVPPQGGPPGFLAHTHAGLQ-----APGFLPAPA 176
Db      173 DLDMERITPALSPCV-----SSSLSEMHMQMDIIPDSTDLVNLQVSPWPSTSEA 224
QY      177 GDKGDLQAVQSCGLADHLITASWGADPVPTKAFGEQEGPLTGACAGGPGLPAGELY 236
Db      225 ATDED-----ERKIAEDLMKLFEOSEWQPTHIDKGY-----LNEPQTQLSSVY 270
QY      237 G 237
Db      271 G 271
```

Search completed: June 18, 2003, 12:48:17  
Job time : 24 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:40:57 ; Search time 38 Seconds  
(without alignments)  
1763.817 Million cell updates/sec

Title: US-09-647-965-9  
Perfect score: 2731  
Sequence: 1 MALAPERAPRVLFGEMILG.....SANSIXDDIECFMELREQPA 503

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
A.Geneseq\_101002:\*

1:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT:*
2:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23:	/SID52/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2731	100.0	503	20	AAV15103
2	2695	98.7	503	22	AA809329
3	1567.5	57.4	542	22	AAV15104
4	458	16.8	427	22	AAW78986
5	454	16.6	427	21	AAV87783
6	450	16.5	427	20	AAV15102
7	446	16.3	175	23	ABP42913
8	376	13.8	450	17	AA899426
9	361.5	13.2	450	17	AA899427
10	357	13.1	451	19	AAW38426

11	335.5	12.3	72	22	ABB11040
12	320	11.7	392	22	ABB11963
13	320	11.7	392	22	AAW79970
14	259.5	9.5	373	19	AAW74784
15	214	7.8	136	21	AA801343
16	198.5	7.3	102	21	AA800953
17	176	6.4	395	23	ABP41797
18	169.5	6.2	329	11	AA805552
19	169.5	6.2	329	11	AA805371
20	167	6.1	1023	23	AAU82954
21	166	6.1	299	11	AA805369
22	166	6.1	325	11	AA805553
23	166	6.1	325	14	AA844219
24	166	6.1	326	11	AA805370
25	166	6.1	329	14	AA844218
26	166	6.1	349	14	AA844217
27	165	6.0	408	22	AAW40720
28	160.5	5.9	551	21	AA842049
29	157	5.7	1413	23	AAU10545
30	155.5	5.7	668	23	AAU10545
31	155.5	5.7	705	22	AAW39328
32	155.5	5.7	714	22	AAW41114
33	155	5.7	561	14	AA837739
34	155	5.7	561	17	AA832449
35	155	5.7	561	19	AAW57650
36	155	5.7	576	21	AA824234
37	155	5.7	576	22	AA895100
38	154.5	5.7	1064	17	AA832554
39	154.5	5.7	1064	19	AAW57652
40	154.5	5.7	1065	14	AA837741
41	154	5.6	1040	22	ABG14734
42	153	5.6	349	22	AA870698
43	153	5.6	349	22	AAU84286
44	153	5.6	542	23	ABB11016
45	152.5	5.6	507	14	AA837682

## ALIGNMENTS

RESULT 1					
ID	AAV15103	standard; Protein; 503 AA.			
XX	AAV15103;				
XX	AC				
XX	25-JAN-2000	(first entry)			
DT	Modified Interferon Regulatory Factor-7 protein (IRF-7).				
DE	Interferon regulatory factor; IRF-7; transactivation domain; serine; threonine; carboxy terminus; aspartic acid; mutant; phosphorylation; post-translational modification; sendai virus; cancer treatment; herpes; activator; promoter; ISRE regulatory element; stimulation; influenza; DNA binding; transcriptional activity; viral infection; HIV infection; activate; homology; cytokine gene; target cell.				
KW					
XX					
OS	Homo sapiens.				
XX	Synthetic.				
XX					
FT	Key	Location/Qualifiers			
FT	Misc-difference 477	/note= "wild type Ser replaced with Asp"			
FT	Misc-difference 479	/note= "wild type Ser replaced with Asp"			
PN	WO9951737-A1.				
XX	14-OCT-1999.				
PD					
XX	07-APR-1999;	99WO-CA00314.			
PF	07-APR-1998;	98CA-2234588.			
XX					
PR					

Human IFN regulato  
Human IFN regulato  
Human protein SEQ  
Human secreted pro  
Human secreted pro  
Human secreted pro  
Human ovarian anti  
Murine interferon  
IRF-1 active prote  
Human homologue of  
Protein coded for  
Human interferon x  
Human interferon x  
IRF-1 active prote  
Murine interferon  
Interferon-beta re  
Human polypeptide  
Human ORF ORF1813  
Human PKIN-24 prot  
Rat synapsin 1B (Y  
Human polypeptide  
Collagen-like poly  
Collagen-like poly  
Collagen-like poly  
Human vesicle asso  
Human protein sequ  
Collagen-like poly  
Collagen-like poly  
Novel human diagno  
Human IRF-2 protei  
Human endometrial  
Human secreted pro  
gIV from BHV-1 str

XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.  
 PA Hiscott J, Lin R;  
 XX MPI, 1999-620201/53.  
 DR N-FSDB; AAX90993.  
 XX  
 PT Carboxy-terminus modified highly active forms of interferon regulatory  
 PT factor proteins used for the treatment of viral infections  
 XX  
 PS Claim 14, Fig 12; 93pp; English.  
 XX  
 CC The present protein sequence is the interferon regulatory factor-7  
 CC (IRF-7) 2D protein, that is modified in the transactivation domain. The  
 CC serine and threonine residues in the carboxy terminus are modified  
 CC post-translationally, by phosphorylation, following sendai virus  
 CC infection. The modified IRF-7, substituted with aspartic acid, functions  
 CC as a activator of promoters containing ISRE regulatory elements and  
 CC stimulation of DNA binding and transcriptional activity. IRF-7 protein  
 CC shows highest homology to IRF-3. The modified IRF proteins are used for  
 CC the treatment of viral infections like, influenza, herpes or HIV  
 CC infection. They may also be used to activate a cytokine gene, in cancer  
 CC treatment or to modify a target cell of an organism.

SQ Sequence 503 AA;

Query Match 100.0%; Score 2731; DB 20; Length 503;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-197;  
 Matches 503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MALAPERAAARVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFAKDLSEADARIF 60  
 1 MALAPERAAARVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFAKDLSEADARIF 60  
 61 KANAVARAGRWPPSSRGCGPPPEAETARACWKTNFCALSTRFVWLKDNSGDPADPHK 120  
 61 KANAVARAGRWPPSSRGCGPPPEAETARACWKTNFCALSTRFVWLKDNSGDPADPHK 120  
 121 VVALSRELCKWRBPGTDTQTEAEPAAVPPGSGPFFLAHTHAGLOAPGLPAPAGDKG 180  
 121 VVALSRELCKWRBPGTDTQTEAEPAAVPPGSGPFFLAHTHAGLOAPGLPAPAGDKG 180  
 181 DILLQAVQOOSCLADHLTLTASWGADVPVTKAPGEGEGPLTGACAGGPGLPAGELYGNV 240  
 181 DILLQAVQOOSCLADHLTLTASWGADVPVTKAPGEGEGPLTGACAGGPGLPAGELYGNV 240  
 241 ETTSPSGPOPALTTGEEAAPSPHQAEPYLSPSACTAVQSPGALDVTIMYKGRIV 300  
 241 ETTSPSGPOPALTTGEEAAPSPHQAEPYLSPSACTAVQSPGALDVTIMYKGRIV 300  
 301 LOKVGHPSCTFLYGPDPVAVRATDPQOVAFPSPALPDOKOLRYTEELRHVAPGLHLE 360  
 301 LOKVGHPSCTFLYGPDPVAVRATDPQOVAFPSPALPDOKOLRYTEELRHVAPGLHLE 360  
 361 LRGPOLMARMGCKKYWEVGGPPGASPSPTACLLPRNCDTPIPFPRVFFOGLVEFRAR 420  
 361 LRGPOLMARMGCKKYWEVGGPPGASPSPTACLLPRNCDTPIPFPRVFFOGLVEFRAR 420  
 421 QRRGSPRYTLYLFGODLSAGRPRKSLVVKLEPMLCRVHLEGTORREGVSSLDSSDLDI 480  
 421 QRRGSPRYTLYLFGODLSAGRPRKSLVVKLEPMLCRVHLEGTORREGVSSLDSSDLDI 480  
 481 CUSANSLYDIDICFLMELEOPA 503  
 481 CUSANSLYDIDICFLMELEOPA 503

RESULT 2  
 AAE09329  
 ID AAE09329 standard; Protein; 503 AA.  
 XX  
 AC AAE09329;

XX 19-NOV-2001 (first entry)  
 DT  
 XX Human intracellular regulatory molecule, KWC02.  
 DE  
 XX Human; intracellular regulator; cell division; proliferation; therapy;  
 KW cancer; infection; wound; developmental abnormality; metabolic problem;  
 KW cytostatic; antibacterial; vulnerary; transcription factor; KWC02;  
 KW interferon response factor; IRF.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6274312-B1.  
 XX  
 PD 14-AUG-2001.  
 XX  
 PF 10-DEC-1997; 97US-0999774.  
 XX  
 PR 11-DEC-1996; 96US-0032818.  
 XX  
 PA (SCHE ) SCHERING CORP.  
 XX  
 PI Gish KC, Seghezzi W, Shanahan F, Lees EM, McClanahan TK;  
 XX  
 DR MPI; 2001-535086/59.  
 DR  
 DR N-PSDB; AAD16105.  
 XX  
 PT New genes encoding intracellular regulatory molecules, useful for  
 PT regulating cell division and proliferation (e.g. tumor cells),  
 PT particularly for treating cancer, infections, wounds, or developmental  
 PT or metabolic abnormalities  
 XX  
 PS Example 2; Column 33-38; 62pp; English.

CC The present sequence is a human intracellular regulatory  
 CC molecule, KWC02. The KWC02, an interferon response factor (IRF)  
 CC homologue, is a transcription factor. The polynucleotides encoding  
 CC intracellular regulatory molecules are useful for regulating cell  
 CC division and proliferation of various cell types, including tumor  
 CC cells. Specifically, they are also useful for treating cancer,  
 CC infections, wounds, developmental abnormalities or metabolic problems.

SQ Sequence 503 AA;

Query Match 98.7%; Score 2695; DB 22; Length 503;  
 Best Local Similarity 99.0%; Pred. No. 8.4e-195;  
 Matches 498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 MALAPERAAARVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFAKDLSEADARIF 60  
 1 MALAPERAAARVLFGEWLLGEISSGCEGLQWLDEARTCFRVPWKHFAKDLSEADARIF 60  
 61 KANAVARAGRWPPSSRGCGPPPEAETARACWKTNFCALSTRFVWLKDNSGDPADPHK 120  
 61 KANAVARAGRWPPSSRGCGPPPEAETARACWKTNFCALSTRFVWLKDNSGDPADPHK 120  
 121 VVALSRELCKWRBPGTDTQTEAEPAAVPPGSGPFFLAHTHAGLOAPGLPAPAGDKG 180  
 121 VVALSRELCKWRBPGTDTQTEAEPAAVPPGSGPFFLAHTHAGLOAPGLPAPAGDKG 180  
 181 DILLQAVQOOSCLADHLTLTASWGADVPVTKAPGEGEGPLTGACAGGPGLPAGELYGNV 240  
 181 DILLQAVQOOSCLADHLTLTASWGADVPVTKAPGEGEGPLTGACAGGPGLPAGELYGNV 240  
 241 ETTSPSGPOPALTTGEEAAPSPHQAEPYLSPSACTAVQSPGALDVTIMYKGRIV 300  
 241 ETTSPSGPOPALTTGEEAAPSPHQAEPYLSPSACTAVQSPGALDVTIMYKGRIV 300  
 301 LOKVGHPSCTFLYGPDPVAVRATDPQOVAFPSPALPDOKOLRYTEELRHVAPGLHLE 360  
 301 LOKVGHPSCTFLYGPDPVAVRATDPQOVAFPSPALPDOKOLRYTEELRHVAPGLHLE 360  
 361 LRGPOLMARMGCKKYWEVGGPPGASPSPTACLLPRNCDTPIPFPRVFFOGLVEFRAR 420

```
Db 361 LRGPLMARMKCKCYWVEVGPPSGASBPSTPACLLPRNCDDPDPDFVFFQELTFEPAR 420
OY 421 ORRGSPRTIYGFQGDLSAGRPKRSVLVLEPMLCRVHLEGNQREGVSLDSSDL 480
Db 421 ORRGSPRTIYGFQGDLSAGRPKRSVLVLEPMLCRVHLEGNQREGVSLDSSDL 480
OY 481 CLSSANSLYDDIECFLEMELEQPA 503
Db 481 CLSSANSLYDDIECFLEMELEQPA 503

RESULT 3
AAV15104
ID AAV15104 standard; Protein; 542 AA.
XX

AAV15104;
25-JAN-2000 (first entry)
Chimeric protein IRF-7(1-246)/IRF-3(132-427).
Chimeric protein IRF-7(1-246)/IRF-3(132-427).
Interferon regulatory factor; IRF; chimeric protein; serine, threonine,
carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;
post-translational modification; sendai virus; cancer treatment; herpes;
PRDI/PRDII; promoter; ISRE regulatory element; stimulation; activate;
activator; DNA binding; transcriptional activity; viral infection;
proteasome mediated degradation; influenza; HIV infection; cytokine gene;
target cell.
Homologous.
Synthetic.
XX
OS
XX
FH Key location/Qualifiers
FT Region 1..246
FT of IRF-7 protein
FT of IRF-7 protein
FT Region 247..542
FT of IRF-3 protein
FT of IRF-3 protein
FT Misc-difference 511
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 513
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 517
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 519
FT /note= "Wild type Ser replaced with Asp"
FT Misc-difference 520
FT /note= "Wild type Ser replaced with Asp"
FT MO9951737-A1.
XX
XX 14-OCT-1999.
XX
XX 07-APR-1999; 99WO-CA00314.
XX
XX 07-APR-1998; 98CA-2234588.
XX
XX (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
XX
XX Hiscott J, Lin R;
XX
XX WPI, 1999-620201/53.
XX
XX N-PSDB; AAV90994.
XX
XX
XX Carboxy-terminus modified highly active forms of interferon regulatory
XX factor protein used for the treatment of viral infections
XX
XX Claim 11; Fig 13; 93pp; English.
XX
XX The present sequence is the chimeric protein comprising, residues 1-246
XX from the modified amino-terminal domain of interferon regulatory factor,
XX
```

```
CC IRF-7 and 132-427 residues from the carboxy-terminus of modified IRF-3
CC (5D) protein. The serine and threonine residues are post-translationally
CC modified by phosphorylation, following sendai virus infection. The
CC modified IRF, substituted with aspartic acid, functions as a strong
CC activator of promoters containing ISRE and PRDI/PRDII regulatory
CC elements, stimulation of DNA binding and transcriptional activity and
CC the proteasome mediated degradation. The modified IRF sequences are used for
CC the treatment of viral infections like, influenza, herpes or HIV
CC infection. They may also be used to activate a cytokine gene, in cancer
CC treatment or to modify a target cell of an organism.
XX
XX SQ Sequence 542 AA;
XX
XX Query Match 57.4%; Score 1567.5; DB 20; Length 542;
XX Best Local Similarity 59.7%; Pred. No. 7.7e-110;
XX Matches 327; Conservative 38; Mismatches 116; Indels 67; Gaps 12;
OY 1 MALAPERAPRVLFGEWMLGFISSGCTGCLQWLEDEARTCFRVPMKHPARKDLSADARIF 60
Db 1 MALAPERAPRVLFGEWMLGFISSGCTGCLQWLEDEARTCFRVPMKHPARKDLSADARIF 60
OY 61 KAMAVARGMPSSRGSGPPPEAETARAGKTNFRCALRSTRFVWLRDSSGDPADPHK 120
Db 61 KAMAVARGMPSSRGSGPPPEAETARAGKTNFRCALRSTRFVWLRDSSGDPADPHK 120
OY 121 VYALSRELCWREGSGTQTEAEAPAVPPGCGPPFPLATHAGLOAPGLPAPADKG 180
Db 121 VYALSRELCWREGSGTQTEAEAPAVPPGCGPPFPLATHAGLOAPGLPAPADKG 180
OY 181 DLLLQAVQSCGLADHLITAGWADVPPTKAPREGQEGPLTGACAGGGLPAGELYGWAV 240
Db 181 DLLLQAVQSCGLADHLITAGWADVPPTKAPREGQEGPLTGACAGGGLPAGELYGWAV 240
OY 241 ETTSS-----PGQPALTTGEAAPSHPQAEPLSPS----- 274
Db 241 ETTSSPTSDTQEDILDELGMVLAPLPDEPSEL---AAVAPPCQ--PLRSPSLDNP 294
OY 275 -----PSACTAVERSPG---ALDVTIMYKGRTVLQKVGHSCFTLYGPPPAVAT 324
Db 275 TTFNLGSENPRLRLVPGSEMEFVTAFRGQVQQTISCEGLRVG-SEVGRRL 353
OY 325 DPQVAPPSPA-ELPDQKRLRYTEELLRHVAPGLHLERGQMLARMKCKCYWVEVG-- 381
Db 325 PGMPVTLDPDPMSTLDGVMGYVHVLSCLGGGLARWAGQMLWAORLGHCHTAVASEE 413
OY 382 -----GPPSASBPSTPACLLPRNCDDPDPDFVFFQELTFEPARQRRGSPRTIYGF 434
Db 382 -----VPRDKSGGVDFLDPFVDDLTFTEGSGR-SPRYALMFCV 463
OY 435 GODLSAGRPKRSVLVLEPMLCRVHLEGNQREGVSLDSSDLCLSSANSI---YDD 491
Db 464 GESWPDQDPWTKRLVYMKVVTCLRALVENARVGGASLENT-VDLHINDHPLDLDQ 522
OY 492 TEEFLMEI 499
Db 523 YKAVIQDL 530

RESULT 4
AAV78986
ID AAV78986 standard; Protein; 427 AA.
XX
XX AAV78986;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 1648.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukemia;
XX nervous system disorder; arthritis; inflammation.
XX
```

OS Homo sapiens.  
 XX  
 PN MO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001MO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
 PI  
 WPI; 2001-476283/51.  
 N-PSDB; AAK52119.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20, Page 3986; 6221PP; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAK7833-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoietic regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAK80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 CC  
 XX  
 XX Sequence 427 AA;  
 XX  
 Query Match 16.8%; Score 458; DB 22; Length 427;  
 Best Local Similarity 28.0%; Pred. No. 2.2e-26;  
 Matches 141; Conservative 64; Mismatches 191; Indels 108; Gaps 19;  
 10 PRLFGEMILGEISSGCEYEGLOLDEARTCFRPMWHPARKDISEADARIFKMAVARGR 69  
 6 PRLT--PWLVSQDLIDGLEGVAMVWNSKRTFRIRIPMKHGRQDQOEDFGIFQAMAEATGA 63  
 70 WPISSRGSGPPPEAETAEARAGWKTFRFVMLRDNSSGDPADPHKVALSRELIC 129  
 64 YV-----GRKPLPFT-----WKNRSLRNKKEGRLIADNRKSDPHDPIKTEFY---- 110  
 130 WRGPGFTDTEAPAPAVPPPGGPPFLAHTHAGIQAAPGPPAPAGDKGDLILQAVQO 189  
 111 -NSGVG---DFSQPTSPPTNG-----GSTSD-----TQ 136  
 190 SCILADHLLTMSKADPPTAPRBEGBGLPLTGACAGGRLPRGELYMAVETTPSPGQ 249  
 137 EDLIDELL--GNWVLAPLP-----DPPPP-----SLAVAPPCPO 169  
 250 PAALTITGEAAAPSPHQAEPYLSPPSACTAVOEP--SPGALDVTIMYKRTVLOKVVGP 308  
 170 PLRSPDLNPTP-----FNLGSPENPLKRLILVPGBEWFEFTATAYRGQVQOITISCP 223  
 309 SCTFLYGPDPDAVRATDPQVAFPSPA-ELPDOKOLRYTEELLRHVAPGLHLELRGPOLM 367

DB 224 EGIRLVIG--SEVGDRTLPGWPTVLPDPGMSLTDRGVMSYVRHVLSCIGGLALMRAGOWLM 282  
 QY 368 ARRMKCKKYWVEVG-----GPPGSASPESTPACLPBRNCDTIPPERVFOQLVER 418  
 DB 283 AQRIGHCHTYMAVASEELLJNSGHPDGE-----VPKDEGVFDLGPVILITFT 333  
 QY 419 ARQRGSPRYTYILFGODLSAGRPKESLYLVKLEPMLCRVHLGTOREGVSSDSSDL 478  
 DB 334 ESSGR--SPRYALMFCVGBSWPDQDPWTKRLVWVKVYPTICRLALVEMARVGAASSLENT-V 391  
 QY 479 DCLSSANSL---YDDIEGFLMEL 499  
 DB 392 DLHISNHPILSLTSDQYKAYLQDL 415  
 RESULT 5  
 AAY87783  
 ID AAY87783 standard; Protein; 427 AA.  
 XX  
 AC AAY87783;  
 XX  
 DT 24-AUG-2000 (first entry)  
 XX  
 DE Human IRF3 protein.  
 XX  
 KM Human; ADA2; cytostatic; gene therapy; treatment; cancer; IRF3.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT Misc-difference 8  
 XX  
 PN US6054289-A.  
 XX  
 PD 25-APR-2000.  
 XX  
 PF 30-AUG-1996; 96US-0705771.  
 XX  
 PR 30-AUG-1995; 95US-0002993.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Moore PA;  
 XX  
 DR WPI; 2000-338491/29.  
 XX  
 DR N-PSDB; AAA39474.  
 XX  
 PT New polynucleotide encoding human AD2 is useful for treating cancer and  
 PT for isolating cDNAs and genes having similar biological activity -  
 XX  
 PS Disclosure; Column 59-62; 54pp; English.  
 XX  
 CC This invention describes a novel polynucleotide (1) encoding human ADA2.  
 CC The products of the invention have cytostatic activity and can be used  
 CC for gene therapy. (1) is useful for treating cancer; as primers and  
 CC probes for isolating full length cDNA and genes having similar  
 CC biological activity. This sequence represents the human IRF3 protein  
 CC described in the method of the invention.  
 CC  
 XX  
 XX Sequence 427 AA;  
 XX  
 Query Match 16.6%; Score 454; DB 21; Length 427;  
 Best Local Similarity 28.0%; Pred. No. 4.3e-26;  
 Matches 141; Conservative 63; Mismatches 192; Indels 108; Gaps 19;  
 10 PRLFGEMILGEISSGCEYEGLOLDEARTCFRPMWHPARKDISEADARIFKMAVARGR 69  
 6 PRLT--PWLVSQDLIDGLEGVAMVWNSKRTFRIRIPMKHGRQDQOEDFGIFQAMAEATGA 63  
 70 WPISSRGSGPPPEAETAEARAGWKTFRFVMLRDNSSGDPADPHKVALSRELIC 129



DT		22-AUG-2002 (first entry)
XX		
DE	Human ovarian antigen HPDRRA44, SEQ ID NO:4045.	
KW	Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;	
KM	ovarian cancer; breast cancer; tumour; reproductive system disorder;	
KM	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;	
KM	PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;	
KM	inflammatory condition; immune disorder; blood disorder;	
KM	cardiovascular disorder; respiratory disorder; neurological disorder;	
KM	gastrointestinal disorder; urinary system disorder; drug screening;	
KM	gene therapy; chromosome mapping; forensic analysis;	
KM	antibody preparation; cytostatic; immunomodulatory; neuroprotective;	
KM	antiinflammatory; gynaecological; reproductive.	
XX		
OS	Homo sapiens.	
PN	WO200200677-A1.	
PD	03-JAN-2002.	
PF	07-JUN-2001; 2001WO-US18569.	
PI	07-JUN-2000; 2000US-209467P.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
PL	Birse CE, Rosen CA;	
DR	WPI; 2002-147876/19.	
XX	N-PSDB, ABQ55990.	
PT	Isolated nucleic acid molecules encoding novel ovarian polypeptides,	
PT	useful in the prevention, treatment and diagnosis of cancer (e.g.	
PT	ovarian cancer), immune disorders, cardiovascular disorders and	
PT	neurological diseases -	
PS	Claim 11; SEQ ID No 4045; 2922pp; English.	
XX	The invention relates to 2175 novel human ovarian antigens (ABP41054-	
CC	ABP41328) and to cDNAs encoding them (ABO54131-ABO56305), and also	
CC	comprehends polypeptides 90% identical and polynucleotides 95% identical	
CC	to the sequences of the invention. The invention additionally relates to	
CC	recombinant vectors and host cells comprising human-ovarian antigen	
CC	polynucleotides, antibodies against human ovarian antigens, and the use	
CC	of ovarian antigen polynucleotides and polypeptides in diagnosing,	
CC	treating, prognosing or preventing various ovary and/or breast-related	
CC	disorders. Such conditions include ovarian cancer and breast cancer, and	
CC	metastatic tumours of ovarian or breast origin, reproductive system	
CC	disorders (e.g., infertility, disorders of pregnancy, anovulation,	
CC	polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine	
CC	disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic	
CC	shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and	
CC	vaginitis), immune disorders (e.g., congenital and acquired	
CC	immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus)	
CC	blood-related disorders (e.g., anaemia), cardiovascular disorders,	
CC	respiratory disorders, neurological disorders, gastrointestinal disorders,	
CC	and urinary system disorders. Ovarian antigen polypeptides and	
CC	polynucleotides may also be used in screening for compounds which	
CC	modulate ovarian antigen expression or activity. The polynucleotides may	
CC	further be used for gene therapy, chromosome mapping, in the	
CC	identification of individuals and in forensic analysis, and the	
CC	polypeptides may be used as food additives or to prepare antibodies	
CC	useful in disease diagnosis, drug targeting and phenotyping. The present	
CC	sequence represents a human ovarian antigen of the invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ffp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 175 AA;	
Query March	16.3%; Score 446; DB 23; Length 175;	
Best Local Similarity	55.1%; Pred. No. 6e-26;	

Matches	102;	Conservative	9;	Mismatches	16;	Indels	58;	Gaps	8;
Oy	334	PAELPDQKQRYTEELLRHVAPGHLRLRGQLMARRMGKKYWEVGGPPGSGASPSSTPA							393
Db	2	PAKLPPDQKQRYTEELLRHVAPGHLRLRGQLMARRMGKKYWEVGGPPGSGASPSSTPA							61
Oy	394	CLLPRNCDTPIFDFRVFFQ-----ELVEFRARQ							421
Db	62	CLLPRNCDTPIFDFRVFFQGGVRLAMGSGKSLGKHLPCQVLAKKSPSELVEF----							117
Oy	422	RRGSPRY-TY---LARGODLSAGRPKEK-----SLVLVKLEPWLCLRVHLEGTGOREG							469
Db	118	RHGAPLWPLTYLHPLGLRAG--FVSLGGPRRRRGPGAGLTVV-----PMLHGTGTHXK-							167
Oy	470	VSSLD 474							
Db	168	VGFLD 172							

[illegible]

Db 26 QWLIDQIDSGKYGLWENEEKSVIRIPWKHAGKODYNNREDDALFKAMALFKGKF---- 81  
 QY 75 RGGGPPPEAEETARAGWKTNFRICALRSTRFRVLMRDNSG--DPADPHKYVALSRELQWREG 133  
 Db 82 REGIDKPPPT-----WKTRLRICALNKSNDDEELVERSDLDISDYKYKIRI----- 127  
 QY 134 PGDTQTEAEAPAAVPPGCGPPG--FLAHTHAGLQAPGLPAPAGDKGDLLLQAVQOS 190  
 Db 128 -----VPEGAKKGAQQLTLDITQMMGHYPMTAP---YGSLPAGQVHNH 169  
 QY 191 CLADHLITASWAGADVPPTKAPGEGEGLPLTGACG-----GGLP-----AGELYGMAYE 241  
 Db 170 MWPRP--DRSW--RDVAPDQSHPEIPYQCPTFGPRGHMWGSPSCNCGCVTGTFFACAP 226  
 QY 242 TPSPG-PPPALTTGEEAAPSPHQAEPYLSPPSACTAVQEPSPGALDVTIMYKGRV 300  
 Db 227 ESQAGGIPLEPSIRSAEALA-----LSDC-----RLHICLYYRDILV 263  
 QY 301 LQKVVGHP--SCTFLYGPDPDAVRATDPQVAPSPAPALPDOKQRYTEELRHVAPGLH 359  
 Db 264 KELTTSPGCRISHG---HTYDVSNLDQVLFPPY---DDNGQRKNIEKLSHLERGLVL 317  
 QY 360 ELRGQPLMARMKGCKKYWEVGGPPGSAS--PSTPACLPKNCCTPIFDPRVFFQELVE 417  
 Db 318 WMAPDGLYAKRLCOSRIYWD--GPLALCSDRENK-----LERDQCKLFDTOQFLSELQVF 371  
 QY 418 RARQRGSPRYTILYGFQDLSAGRPKEKSLVVKLEPWLQR 459  
 Db 372 -AHHGRPAFRFOVTLQGEPEPDPQ--RQRKLITAHVEPLLAR 411

RESULT 9  
 AAR99427  
 ID AAR99427 standard; Protein; 450 AA.

AC AAR99427;  
 DT 20-JAN-1997 (first entry)  
 XX Human lymphocyte specific interferon regulatory factor.  
 DE Human lymphocyte specific interferon regulatory factor.  
 KW Lymphocyte specific interferon regulatory factor; LSIRF; IRF-3.  
 XX Homo sapiens.  
 OS

Key Location/Qualifiers  
 Misc-difference 163..164  
 FT /note= "the Double Q form of hLSIRF contains an additional glutamine residue between residues 163 and 164 of the single Q form"

FT WO9632477-A1.  
 XX 17-OCT-1996.  
 PD 12-APR-1996; 96WO-CA00231.  
 PF 03-APR-1996; 96US-0611280.  
 PR 14-APR-1995; 95US-0422733.  
 XX (AMGE-) AMGEN CANADA INC.  
 PA Grossman A, Matsuyama T, Richardson CD;  
 XX WPI; 1996-477128/47.  
 DR N-PSDB; AAT411706.  
 XX New genes for murine lymphocyte specific interferon regulatory  
 PT factor - used for modulation of lymphocyte activation and  
 PT proliferation  
 XX Claim 1; Page 65-66; 92pp; English.  
 PS  
 XX

CC Novel single Q or Double Q forms of human lymphocyte-specific  
 CC interferon regulatory factor (LSIRF or ISF-3) (AAR99426) bind to  
 CC interferon-stimulated response elements in DNA. The amino acid  
 CC sequence of the single Q form was deduced from a cDNA clone (AAT41706)  
 CC isolated from a human lymphocyte cDNA library. Recombinant LSIRF  
 CC can be produced in transformed prokaryotic or eukaryotic host  
 CC cells. LSIRF polypeptides are useful as targets for therapeutic  
 CC cpds. used to regulate lymphocyte activation. By blocking LSIRF  
 CC activity it is possible to suppress lymphocyte activation in  
 CC response to certain environmental stimuli.  
 XX SQ Sequence 450 AA:

Query Match 13.2%; Score 361.5; DB 17; Length 450;  
 Best Local Similarity 27.2%; Pred. No. 4,36-19;  
 Matches 135; Conservative 70; Mismatches 194; Indels 97; Gaps 22;

QY 16 EMLGEISSGCEGLQWIDEARTCFRVWPKHARKDIS--EADARIFKAMAVARGRWPSS 74  
 Db 26 QWLIDQIDSGKYGLWENEEKSVIRIPWKHAGKODYNNREDDALFKAMALFKGKF---- 81  
 QY 75 RGGGPPPEAEETARAGWKTNFRICALRSTRFRVLMRDNSG--DPADPHKYVALSRELQWREG 133  
 Db 82 REGIDKPPPT-----WKTRLRICALNKSNDDEELVERSDLDISDYKYKIRI----- 134  
 QY 134 PGDTQTEAEAP--AAVPPGCGPPGPF--AHTHAGLQAPGLPAPAGDKGDLLLQAVQOS 190  
 Db 135 -GAQQLTLEDQMSMSHYTMTTPYSLPAGVHNHMP----- 172  
 QY 191 CLADHLITASWAGADVPPTKAPGEGEGLPLT-----GACAGGGLPAGELYGMAY 240  
 Db 173 -----LDRSW--RDVVPDQHPRELIPYQCPTFGPRGHMWGSPSCNCGCVTGTFFACAP 224  
 QY 241 ETTPSPG-PPPALTTGEEAAPSPHQAEPYLSPPSACTAVQEPSPGALDVTIMYKGRV 299  
 Db 225 PESQAPGVTEPSIRSAEALA-----FSDC-----RLHICLYYREIL 261  
 QY 300 LQKVVGHP--SCTFLYGPDPDAVRATDPQVAPSPAPALPDOKQRYTEELRHVAPGLH 358  
 Db 262 VKELTSPGCRISHG---HTYDVSNLDQVLFPPY---ENNGQRKNIEKLSHLERGLV 315  
 QY 359 LELRGQPLMARMKGCKKYWEVGGPPGSAS--PSTPACLPKNCCTPIFDPRVFFQELVE 416  
 Db 316 WMAPDGLYAKRLCOSRIYWD--GPLALCNDRENK-----LERDQCKLFDTOQFLSELQ 369  
 QY 417 FRARQRGSPRYTILYGFQDLSAGRPKEKSLVVKLEPWLCRYHLESTQRECVSLDSS 476  
 Db 370 F-AHHGRSLPFRFOVTLQGEPEPDPQ--RQRKLITAHVEPLARQLYFPAQONSQHFLRGY 427  
 QY 477 DLDCLSGANSLYDDI 492  
 Db 428 DLPEHISNPEYHRSI 443

RESULT 10  
 AAW38426  
 ID AAW38426 standard; Protein; 451 AA.  
 AC AAW38426;  
 DT 08-JUN-1998 (first entry)  
 XX Human multiple myeloma oncogene 1 (MM-1) product.  
 DE Multiple myeloma oncogene 1; MM1 protein; tumour; human;  
 KW diagnosis; interferon regulatory factor 4; IRF4.  
 XX Homo sapiens.  
 OS  
 XX WO9745106-A1.  
 XX 04-DEC-1997.  
 XX







QY 360 ELRGPOLMARBMCKKYWEVG-----GPPGSASPTPACLLPRNCTPIFDPRVF 410  
 DB 234 WRAGOMLMAQRIGHCHTYMAVSEELLPNSGCHPDGE-----VPKDEGGVFDLGP 284  
 QY 411 F-----QELVEFRARQRGSPRYTYLGFQDLSAGRPKESLYLVLEPMLCRVHLEG 464  
 DB 285 IVSGSLPPDLITFTESGR-SPRYALMFCVGSWPDDQPMTKLVVVKVPTCLRALVEM 343  
 QY 465 TORGVSSLDSDLDLCLSSANSL---YDDIECFMEL 499  
 DB 344 ARVGASASLENT-VDLHISNSHPLSLTSDQYKAYLODL 380

RESULT 13  
 AAM79970  
 ID AAM79970 standard; Protein; 392 AA.  
 AC AAM79970;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 Human protein SEQ ID NO 3616.  
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KM nervous system disorder; arthritis; inflammation.  
 OS Homo sapiens.  
 XX  
 PN WO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 PR 20-JUN-2000; 2000US-0598075.  
 PR 19-JUL-2000; 2000US-0620325.  
 PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 DR WPI; 2001-476283/51.  
 N-PSDB; AAK53103.  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 PT useful in diagnosis and gene therapy -  
 XX  
 PS Claim 20; Page 397; 6221pp; English.  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK60302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM60020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.  
 XX  
 SQ Sequence 392 AA.  
 Query Match 11.7%; Score 320; DB 22; Length 392;  
 Best Local Similarity 26.0%; Pred. No. 4.8e-16;  
 Matches 119; Conservative 51; Mismatches 176; Indels 112; Gaps 19;

QY 62 AAVARGMPPSPRGGGPPPEAETERAGWKTNFRCAIARRRPFMYALDNSGDPADPHKV 121  
 DB 15 AAEATGAVVP---GRKPDLPFT---WKNFRSALNRKGLRLADRSKDPDHPKI 65  
 QY 122 YALSRELCEBEGPTDQTEAEAPAAVPPGSGPPFLAHTHAGIQAPGPIAPAGDQCD 181  
 DB 66 YEFV-----NSGVG---DFSGPDISPDTNG-----GTSND 93  
 QY 182 ILLQAVQSCIADHLLTASWGADPVPTYAPGBGQGLPLTGACAGGGLPAGELYGMAVE 241  
 DB 94 -----TQEDILDEL-GNMYLAPLP-----DPGPP-----SLA 120  
 QY 242 TTPSPQPPAALTGGAAPSPHQAEPYLSPSACTAVQEP-SFGALDVTIMYKGRTV 300  
 DB 121 VAEPCPPPLRSPSLDNPTP-----FPNLGSENPRLRLVPGSEWEFVTAFFRGQV 174  
 QY 301 LQKVGHSPCTFLYPPDPVAVATDPQOVAFPSPA-ELPDQKOLRYTEELLRHVAPGLH 359  
 DB 175 FQGTISCEBGLRVG-SEVGDRITLPGWPTLIPDPMSLTRGQVSVYHHVILSCLOGGLAL 233  
 QY 360 ELRGPOLMARBMCKKYWEVG-----GPPGSASPTPACLLPRNCTPIFDPRVF 410  
 DB 234 WRAGOMLMAQRIGHCHTYMAVSEELLPNSGCHPDGE-----VPKDEGGVFDLGP 284  
 QY 411 F-----QELVEFRARQRGSPRYTYLGFQDLSAGRPKESLYLVLEPMLCRVHLEG 464  
 DB 285 IVSGSLPPDLITFTESGR-SPRYALMFCVGSWPDDQPMTKLVVVKVPTCLRALVEM 343  
 QY 465 TORGVSSLDSDLDLCLSSANSL---YDDIECFMEL 499  
 DB 344 ARVGASASLENT-VDLHISNSHPLSLTSDQYKAYLODL 380

RESULT 14  
 AAM74784  
 ID AAM74784 standard; Protein; 373 AA.  
 AC AAM74784;  
 XX  
 DT 25-JAN-1999 (first entry)  
 XX  
 DE Human secreted protein encoded by gene 55 clone HPRC93.  
 KM Human; secreted protein; testis; tumour; foetal brain tissue;  
 KM fusion protein; cancer; central nervous system; seizure;  
 KM diagnosis; neurodegenerative disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 175  
 FT label= unknown  
 FT Misc-difference 373  
 FT label= unknown  
 PN WO9839448-A2.  
 PD 11-SEP-1998.  
 XX  
 PF 06-MAR-1998; 98WO-US04493.  
 PR 02-OCT-1997; 97US-0061060.  
 PR 07-MAR-1997; 97US-0038621.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040162.



QY 291 VTIWYGRVTLQKV-VGHP-SCFTLYGPPDPAPVRAID-----PQVAFPSPAELPDQX 342  
 Db 133 IKFOYRKGKEXGQMTVSNPQGCRLFYGDLPMPDQDELFGPVXLEQVKFPGPHITNEKQ 192  
 QY 343 LRTTEBLRVAAGLHLELRGPQMARMKCKVYEVGSPGASASTACLLPNCDT 402  
 Db 193 KLFSTSLDLVMDRGLILEVSGHAIVAIRLQCCKVYVSGPCAFSLVAVN-----LIEQKRV 248  
 QY 403 PIFDFRVFQELVEFPARQSGSPRYTILYFGODLSAGRPKEKSLVVLKLEPMLCRVHL 462  
 Db 249 KLFCLFETFLSDLIARHKGQIEKOPPEIYLCFGEFMPDGKPLERKILIVQIVAVAMTY 308  
 QY 463 EGTQREGVSSLDSSDLCLSSANSLYDIECFIMEI 499  
 Db 309 EMFSGDFTSRFSDSGSVRLQISTPD-IKDNIWAQLKQL 344

RESULT 15  
 ID AAG01343 standard; Protein; 136 AA.  
 XX AAG01343;  
 AC AAG01343;

06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 5424.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping.  
 XX Homo sapiens.

OS Homo sapiens.  
 XX EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST ) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR MPI; 2000-500381/45.

XX DR N-PSDB; AAC01349.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX

XX Claim 13; SEQ ID 5424; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 136 AA;

Query Match 7.8%; Score 214; DB 21; Length 136;  
 Best Local Similarity 36.3%; Pred. No.1.3e-08;  
 Matches 45; Conservative 21; Mismatches 46; Indels 12; Gaps 3;

QY 17 WLIGELSSGCEYGLQWLDEARTCFRVPMWGHFARKDLS-ADARIFKAWAVARGRPPSSR 75  
 Db 15 WYVEQVESQCFPGVCWDDTAKTMFRLPMWGHAGKQDFREDQDAAPFAVAIFGKY----- 69  
 QY 76 GGGPPEAETARAGKKTFRFCAALSTRFFVWLBDNSG-DPADPHKYVALSRELCEBGP 134  
 Db 70 -----KEGDTGPAVWKTRLCALNKSSEFKEVPERGRMDVAPYKVOQLPBGIVSXQP 124  
 QY 135 GTDQ 138  
 Db 125 GTQK 128

Search completed: June 18, 2003, 12:46:45  
 Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 12:42:52 ; Search time 15 Seconds

(Without alignments)  
1390.840 Million cell updates/sec

Title: US-09-647-965-9

Perfect score: 2731

Sequence: 1 MALAPERRAARVLFGEWLIG.....SANSLYDIECFIMEHQPA 503

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40.\*

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	2715	99.4	503	IRF7_HUMAN	Q92965 homo sapien
2	1562	57.2	457	IRF7_MOUSE	P70434 mus musculu
3	780	28.6	491	IRF3_CHICK	Q90643 gallus gall
4	487.5	17.9	498	IRF5_HUMAN	O13568 homo sapien
5	458	16.8	427	IRF3_HUMAN	O14653 homo sapien
6	444.5	16.3	497	IRF5_MOUSE	P56477 mus musculu
7	438	16.0	419	IRF5_MOUSE	P70671 mus musculu
8	436	16.0	467	IRF6_MOUSE	P97431 mus musculu
9	431	15.8	467	IRF6_HUMAN	O14896 homo sapien
10	376	13.8	450	IRF4_MOUSE	O64287 mus musculu
11	363	13.3	451	IRF4_HUMAN	O15106 homo sapien
12	338	12.4	425	ICSB_CHICK	Q90871 gallus gall
13	333.5	12.2	424	ICSB_MOUSE	P23611 mus musculu
14	328.5	12.0	333	IRTF_HUMAN	Q00978 homo sapien
15	308.5	11.3	426	ICSB_HUMAN	Q02556 homo sapien
16	259.5	9.5	399	IRTF_MOUSE	O61179 mus musculu
17	176	6.1	325	IRF1_HUMAN	P10914 homo sapien
18	167	6.1	349	IRF2_HUMAN	P14316 homo sapien
19	167	6.1	1083	T2D3_HUMAN	O00268 homo sapien
20	166	6.1	329	IRF1_MOUSE	P15314 mus musculu
21	166	6.1	349	IRF2_MOUSE	P23906 mus musculu
22	165.5	6.1	328	IRF2_PAT	P23570 rattus norv
23	164	6.0	348	IRF2_CHICK	O98925 gallus gall
24	163	6.0	333	IRF1_CHICK	Q90876 gallus gall
25	159.5	5.8	429	XPN2_XENLA	P17437 xenopus lae
26	157.5	5.8	705	IRF2_HUMAN	P17600 homo sapien
27	156	5.7	1888	CAIE_CHICK	P32018 gallus gall
28	156	5.7	3149	TEGU_EBV	P03186 Epstein-bar
29	155	5.7	670	SYN1_MOUSE	O88935 mus musculu
30	154	5.6	660	YH11_EBV	P03181 Epstein-bar
31	154	5.6	3530	MY15_HUMAN	Q9UK77 homo sapien
32	152.5	5.6	1464	CA11_MOUSE	P08121 mus musculu
33	150	5.5	704	SYN1_RAT	P09951 rattus norv

34	148.5	5.4	560	MIS_HUMAN	P03971 homo sapien
35	148	5.4	1464	CA11_HUMAN	P02452 homo sapien
36	148	5.4	1466	CA13_HUMAN	P02461 homo sapien
37	146.5	5.4	1453	CA11_CHICK	P02457 gallus gall
38	145.5	5.3	1460	CA11_CANFA	O9x517 canis fami
39	145.5	5.3	1516	CA1H_HUMAN	P39060 homo sapien
40	145	5.3	296	CCO1_CAEEL	P08124 caenorhabd
41	144.5	5.3	1527	CA2B_MOUSE	P39061 mus musculu
42	144.5	5.3	1736	CA2B_HUMAN	P13942 homo sapien
43	144	5.3	415	SYN1_CANFA	O62732 canis fami
44	144	5.3	503	CA19_CHICK	P12106 gallus fami
45	144	5.3	706	SYN1_BOVIN	P17599 bos taurus

## ALIGNMENTS

RESULT 1  
IRF7\_HUMAN STANDARD; PRT; 503 AA.  
AC Q92965; O00331; O00332; O00333; O75924;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Interferon regulatory factor 7 (IRF-7).  
GN IRF7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RC TISSUE=Splice;  
RA Grossman A., Nichol J., Antonio L., Luehly R., Suggs S.,  
RA Sutherland G.R., Mak T.W.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
RX MEDLINE=97459673; PubMed=9315633;  
RA Zhang L., Pagano J.S.;  
RT "IRF-7, a new interferon regulatory factor associated with Epstein-  
RL Barts virus latency.";  
RN Mol. Cell. Biol. 17:5748-5757(1997).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM D).  
RX MEDLINE=99003279; PubMed=9786932;  
RA Au W.-C., Moore P.A., Lafleur D.W., Tombal B., Pitha P.M.;  
RT "Characterization of the interferon regulatory factor-7 and its  
RT potential role in the transcription activation of interferon A  
RT gene.";  
RN J. Biol. Chem. 273:29210-29217(1998).  
RL J.  
CC - FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-  
CC STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q  
CC PROMOTOR (OP) OF EBV NUCLEAR ANTIGEN-1 (EBNA1).  
CC - SUBCELLULAR LOCATION: Nuclear (Potential).  
CC - ALTERNATIVE PRODUCTS: 4 ISOFORMS: A (SHOWN HERE), B/BETA, C/GAMMA  
CC AND D/H; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC - TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN SPLEEN, THYMUS, AND  
CC PERIPHERAL BLOOD LEUCOCYTES.  
CC - SIMILARITY: BELONGS TO THE IRF FAMILY.  
CC -  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; U73036; AAB17190.1; -  
CC EMBL; U53830; AAB80686.1; -  
CC EMBL; U53831; AAB80688.1; -  
CC EMBL; U53832; AAB80690.1; -

DR EMBL; AF076494; AAC70999.1; --  
 DR HSSP; P23906; 2IRF.  
 DR TRANSFAC; T04674; --  
 DR Genew; HGNC:6122; IRF7.  
 DR MIM; 605047; --  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF; 1.  
 DR PRINTS; PR00267; INTERNEGCT.  
 DR PRODOM; PD002355; IRF; 1.  
 DR SMART; SM00348; IRF; 1.  
 DR PROSITE; PS00601; IRF; 1.  
 DR Transcription regulation; DNA-binding; Nuclear protein; Activator;  
 KW Alternative splicing.  
 FT VASPLIC 13 122  
 FT VASPLIC 1 6  
 FT VASPLIC 152 164  
 FT VASPLIC 165 503  
 FT VASPLIC 228 256  
 FT CONFLICT 179 179  
 FT CONFLICT 412 412  
 FT SEQUENCE 503 AA; 54278 MW; AA6A39E0E272727C CRC64;  
 Query Match 99.4%; Score 2715; DB 1; Length 503;  
 Best Local Similarity 99.4%; Pred. No. 3.8e-148;  
 Matches 500; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 MALPERAPRVLTGEMWLTGSSCYEGLOMLDEARTCFVPMWGHFARKDLSEADARIF 60  
 1 MALPERAPRVLTGEMWLTGSSCYEGLOMLDEARTCFVPMWGHFARKDLSEADARIF 60  
 61 KAMAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 120  
 61 KAMAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 120  
 61 KAMAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 120  
 61 KAMAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 120  
 121 VYALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 180  
 121 VYALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 180  
 121 VYALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 180  
 121 VYALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 180  
 181 DLILOAVQOGLADHLLTASWGADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 240  
 181 DLILOAVQOGLADHLLTASWGADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 240  
 181 DLILOAVQOGLADHLLTASWGADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 240  
 181 DLILOAVQOGLADHLLTASWGADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 240  
 241 ETTSPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGRIV 300  
 241 ETTSPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGRIV 300  
 241 ETTSPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGRIV 300  
 241 ETTSPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGRIV 300  
 301 LOKVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLHLE 360  
 301 LOKVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLHLE 360  
 301 LOKVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLHLE 360  
 301 LOKVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLHLE 360  
 361 LRGPOLMARWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 420  
 361 LRGPOLMARWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 420  
 361 LRGPOLMARWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 420  
 361 LRGPOLMARWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 420  
 421 QRRGSPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 480  
 421 QRRGSPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 480  
 421 QRRGSPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 480  
 421 QRRGSPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 480  
 481 CLSSANSIYDIECEFLMELEOPA 503  
 481 CLSSANSIYDIECEFLMELEOPA 503  
 481 CLSSANSIYDIECEFLMELEOPA 503  
 481 CLSSANSIYDIECEFLMELEOPA 503

RESULT 2  
 IRF7\_MOUSE STANDARD; PRT; 457 AA.  
 AC P70434;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Interferon regulatory factor 7 (IRF-7).

GN IRF7.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Spleen;  
 RA Grossman A., Nicholl J., Antonio L., Luethy R., Suggs S.,  
 RA Sutherland G.R., Mak T.W.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-  
 CC STIMULATED RESPONSE ELEMENT (ISRE) IN IFN PROMOTERS AND IN THE Q  
 CC PROMOTER (Q) OF EBV NUCLEAR ANTIGEN-1 (EBNA1) (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U73037; AAB18626.1; --  
 DR HSSP; P23906; 1IRF.  
 DR MGI; MGI:1859212; Ir7f.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF; 1.  
 DR PRINTS; PR00267; INTERNEGCT.  
 DR PRODOM; PD002355; IRF; 1.  
 DR SMART; SM00348; IRF; 1.  
 DR PROSITE; PS00601; IRF; 1.  
 KW Transcription regulation; DNA-binding; Nuclear protein; Activator.  
 FT DNA BIND 11 122  
 FT SEQUENCE 457 AA; 51222 MW; 30B102C68F56142 CRC64;  
 Query Match 57.2%; Score 1562; DB 1; Length 457;  
 Best Local Similarity 62.6%; Pred. No. 1.7e-82;  
 Matches 315; Conservative 36; Mismatches 100; Indels 52; Gaps 7;

3 LABERAPRVLTGEMWLTGSSCYEGLOMLDEARTCFVPMWGHFARKDLSEADARIF 62  
 1 MAEVRGQVRLFGDWLTGESSQYEGLOMLNARTFVPMWGHFARKDLSEADARIF 60  
 63 WAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 119  
 61 WAVARGWPPSSSGGPPPEAEETAEERAKWTFRCALRSTRFVMDLNSGDPADPHK 119  
 120 KYVALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 179  
 120 KYVALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 179  
 120 KYVALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 179  
 120 KYVALSRELCWREGGCTDTEAEPAPAVPPPOGGPPFLATHAGLOAGPLAPADPHK 179  
 180 GDILLQVLOYS---HILSESADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 238  
 178 GDILLQVLOYS---HILSESADVPPTKAPGEGEGPLTGACAGGGLPAGELYGAV 220  
 239 AVETTSPPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGR 298  
 239 AVETTSPPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGR 298  
 239 AVETTSPPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGR 298  
 239 AVETTSPPGPPALTTGGAAPESPHQAEPIYSPSACTAVOEPSGALDVTIMYKGR 298  
 221 QVEAVPSPPQPPALT-----EESLGLDVTIMYKGR 252  
 221 QVEAVPSPPQPPALT-----EESLGLDVTIMYKGR 252  
 221 QVEAVPSPPQPPALT-----EESLGLDVTIMYKGR 252  
 221 QVEAVPSPPQPPALT-----EESLGLDVTIMYKGR 252  
 299 TVLOKVVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLH 358  
 299 TVLOKVVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLH 358  
 299 TVLOKVVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLH 358  
 299 TVLOKVVGHPSCTFLYGGPPPAVATDPOVAFPSPAELPDQKOLRYTEELLRHVAPGLH 358  
 313 LELRGPSIMALRMWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 372  
 313 LELRGPSIMALRMWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 372  
 313 LELRGPSIMALRMWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 372  
 313 LELRGPSIMALRMWKKCKVYWEVGGPGSASPTPACLLPNCCTPIFDPRVFQELVEPPAR 372  
 419 ARORRGSPPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 478  
 419 ARORRGSPPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 478  
 419 ARORRGSPPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 478  
 419 ARORRGSPPRTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 478  
 373 ARRRGSPHTIYLGEGODLSAGRPEKSLVLYKLEPWLGRVHLEGTORGVSLSDDL 432

QY 479 DLCLASSANSYDDIECFLEMLEEQ 501  
 DB 433 GLCSTSTNSLYEDIHFHMDLQ 455

RESULT 3  
 ID IRF3 CHICK STANDARD; PRT; 491 AA.

AC 090643;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE Interferon regulatory factor 3 (IRF-3).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

NCBI\_TaxID=9031;

SEQUENCE FROM N.A.

RA MEDLINE=95334365; PubMed=7541908;  
 Grant C.E., Vasa M.Z., Deeley R.G.;  
 "cIRF-3, a new member of the interferon regulatory factor (IRF) family that is rapidly and transiently induced by dsRNA."  
 Nucleic Acids Res. 23:2137-2146(1995).  
 CC -1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U20338; AAA86995.1; --  
 HSSP; P15314; 1IF1.  
 InterPro: IPR001346; IRF.  
 Pfam: PF00605; IRF.1.  
 PRINTS; PR00267; INTERPROGCT.

DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 KM Transcription regulation; DNA-binding; Nuclear protein; Activator.  
 FT DNA\_BIND 14 112 TRYPTOPHAN PENTAD REPEAT.  
 SQ SEQUENCE 491 AA; 54441 MW; CAE0C2AA8B8376D9 CRC64;

Query Match 28.6%; Score 780; DB 1; Length 491;

Best Local Similarity 38.5%; Pred. No. 6.6e-38;  
 Matches 208; Conservative 60; Mismatches 180; Indels 92; Gaps 17;

QY 2 ALAPARAPVLTGFWLLGSISSGCEYGLQWLDARTCFVPMKGFARKQLSEADARIEK 61  
 DB 3 ALPSEGAQKLRGCPMLNNAVSSGLYGLCWIPDRIRFIRIPKHNARKDVTSSDVEIFK 62  
 QY 62 AMAVARGWRPSSSGGPPPEAETAEBAAGKTKPRCLRSTRFVVMKRDNSGPPADPHKY 121  
 DB 63 AMAKASGRY-----EGNAEDPAKMKTKPRCLRSTRFVVMKRDNSGPPADPHKY 111  
 QY 122 YALSRELQWREGPDTQTEAEPAPVPPQ--GGP-----GEPFLATHAGLQ---- 167  
 DB 112 YAVA-----SGVNDRGSGGFPVAGALQOQPOLLNHNHDLALNTPT 152  
 QY 168 -AGPELPAPAGDKDL-LLOAVQOSCLADHLHLTASGADVPVTPYABEG--QSGLPITG- 222  
 DB 153 DSTGVAAALTVQDLDLQSVLQHCHNISAL-----GSGPTLVAHTGDALPEDALLPQ 207

QY 223 -ACAGGPGLPAGELYGW-AVETTPSRGPQP-----AALTGGAAPESPHQAP 269  
 DB 208 DGLRPE-----QFQWRQLBEPRLNGNPLRGVGGCGDGAALPVPSECALPAPSPAE 262  
 QY 270 YL---SPSPSACTAVQEPSGALDVTIMYKGRITVQKYVHPSCFTLYGPPDPAVRATD 325  
 DB 263 LTFQSANPAPPPAGDIGLPPLDITITIRYRGKMYVQEQVDSRCVLAVQPLDPV--AE 320  
 QY 326 PQGVAFSPAPLPDQKALRYTEELLRHVAPGHLRLRGPLVARMKGRKMYV----- 380  
 DB 321 QRVLVFPSPASLPDPQRRTEDLE--VAGRLERPAQQLATRLKCKVFWALSQGLE 378  
 QY 381 GPPGSGASPTPACLPKNCPTPIPDFRVFOELVEFRARORGSPPRYTLYGCDLSA 440  
 DB 379 GSEP-----PLNLHROQETITPDFRVFCHELRNFRSRKRSRSDFTIFLCGCFSS 431  
 QY 441 GRPKESLVLVLEPMLCRVHLEGTQREGVSSLDSSDLCLCLASSANSYDDIECFLEMLE 500  
 DB 432 TRPKESKLITVLVLPQFCSEYWEQVQRGGASSLNSGNVSLQSLDSFNLFELEQYHMQTD 491

RESULT 4

ID IRF5 HUMAN STANDARD; PRT; 498 AA.

AC Q13568; Q9BQF0;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Interferon regulatory factor 5 (IRF-5).  
 GN IRF5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;  
 RA Grosseman A., Mitrucker H.W., Lantoni L., Mak T.W.;  
 Submitted (APR-1996) to the EMBL/Genbank/DDB databases.  
 RN [2]  
 CC SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 Submitted (MAR-2001) to the EMBL/Genbank/DDB databases.

CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U51127; AAA96056.1; --  
 CC EMBL; BC004201; AA04201.1; --  
 CC EMBL; BC004139; AA04139.1; --  
 CC HSSP; P23906; 1IRF.  
 CC Genew; HGNC:6120; IRF5.  
 CC InterPro: IPR001346; IRF.  
 CC Pfam: PF00605; IRF.1.  
 CC PRINTS; PR00267; INTERPROGCT.

CC PRODOM; PD002355; IRF.1.  
 CC SMART; SM00348; IRF.1.  
 CC PROSITE; PS00601; IRF.1.  
 CC Transcription regulation; DNA-binding; Nuclear protein.  
 FT DNA\_BIND 16 118 TRYPTOPHAN PENTAD REPEAT.  
 FT DOMAN 142 149 POLY-GLU  
 FT CONFLICT 161 175 EDVKNPPTIQPTLR --> DAVQSGPHMTPTSLKEDVKN  
 FT (IN REF. 1).  
 SQ SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

Query Match 17.9%; Score 487.5; DB 1; Length 498;  
 Best Local Similarity 28.9%; Pred. No. 3.1e-21;  
 Matches 154; Conservative 59; Mismatches 173; Indels 147; Gaps 21;

QY 1 MALAPERAAPVLFGEWLGIEISSGCEGLWLDARTCFVPMKFAKDL-SADARI 59  
 DB 7 VAPPPPR---RVRLKPMVLAQVNSQYGLQWVNGEKLFCIPRHATHRHSPDGDWTI 63  
 QY 60 FKAAVAVAGRPSPSRGGCPPEAEATARAGMKTNRCAALSTRFVWLRLNSGD-PADP 118  
 DB 64 FKAAKETGKYTEG-----VDEADPAKKAALRCALKSHDFRLITYGPRDMPQP 114  
 QY 119 HKVALSLREL-----WRGPGTDQTEAA-----PAVP 148  
 DB 115 YKIV---EVCNGAPPTDSQPRPDYSRGAEEBEEBELQRLMPLSLTDEVKMPRLQ 170  
 QY 149 PPGCGPGPFLAHTHAGLQAP-----GPLPAPAGKGLLLQAVQSCDLADHLTAS 200  
 DB 171 PPTLRPP-----TLQPTLPQPVVLGP-PAP-----195  
 QY 201 WGADEPVPFKARGEQEGPLTGACAGGRLAG--ELYGMAVENTTPSRGPALTTGEA 258  
 DB 196 ---DPSPLAPP-----PGNPAQRELLSEVLE---FGLPABL-----227  
 QY 259 AAPSPHQAEPYLSPPSPACTAVOEPSPGALDVTIMYKRTVLQKVGH-PSCFLYGP 317  
 DB 228 --PAGBGLPDLILSPMLLTD-----LEIKQYGRPRPALITINPGCHLFTS-- 277  
 QY 318 DPAVRAIDPQ-----QVAPSPAELPDQQLRTYBELLRHVAFGLHLELGPOLW 367  
 DB 278 --QEAATQEVLELPGPISLEQVRFSPEDIPSDKORFYTNOILLDGLILLQLOGQDLY 335  
 QY 368 ARMGKCKVYVEGSPGSPASPTACLLPNCOTPIPDVFPQGVLEVPARBRSGPR 427  
 DB 336 AIRLCQCKVFW--SGPCASADSCPN-PIQREVKTKLSLHFLNELLLFGKQNTNPP 392  
 QY 428 YTIYVFGQDLSAGRPKESLVLKLEPWLGRVHLEGTORGVSLSDDLDL 480  
 DB 393 FEIFCFGEEMPRDKRPREKULITVQVAVARLLLEMSGELSADSIQLQI 445

## RESULT 5

IRF3 HUMAN STANDARD; PRT; 427 AA.  
 AC 014653:  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Interferon regulatory factor 3 (IRF-3).  
 GN IRF3.  
 OS Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_TaxID=9606;  
 CX (1)  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Retina;  
 RC MEDLINE=96102173; PubMed=8524823;  
 RA Au W.W.-C., Moore P.P.A., Lowther W.W., Huang Y.-T., Pitha P.M.;  
 RT "Identification of a member of the interferon-stimulated response element and  
 RT activates expression of interferon-induced genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:11657-11661 (1995).  
 RN (2)  
 RP SEQUENCE OF 323-413 FROM N.A.  
 RX MEDLINE=99020108; PubMed=9803267;  
 RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;  
 RT "Mapping of human interferon regulatory factor 3 (IRF3) to chromosome  
 RT 19q13.3-13.4 by an intergenic polymorphic marker";  
 RL Ann. Hum. Genet. 62:231-234 (1998).  
 CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH  
 CC OTHER TRANSCRIPTIONAL FACTORS. POSSIBLY MEMBERS OF THE STAT  
 CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT

CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED CONSTITUTIVELY IN A VARIETY OF  
 CC TISSUES.  
 CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z56281; CA91227.1; -;  
 DR EMBL; U86636; AAC6818.1; -;  
 DR HSSP; P23906; 21RF.  
 DR TRNSPAC; T04673; -;  
 DR Genew; HGNC:6118; IRF3.  
 DR MIR; 603734; -;  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERNSPECT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 KW Transcription regulation; DNA-binding; Nuclear protein; Activator;  
 KW Polymorphism.  
 FT DNAS\_BIND 7 107 TRYPTOPHAN PENTAD REPEAT.  
 FT VARIANT 96 96 R->Q (IN DBSNP:968457).  
 FT VARIANT 377 377 /FtId=VAR\_011901.  
 FT VARIANT 427 427 E->K (IN DBSNP:1049486).  
 FT VARIANT 427 427 /FtId=VAR\_011902.  
 FT FT S->T (IN DBSNP:7251).  
 FT FT /FtId=VAR\_011903; CRC64;  
 SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;

Query Match 16.8%; Score 458; DB 1; Length 427;  
 Best Local Similarity 28.0%; Pred. No. 1.3e-19;  
 Matches 141; Conservative 64; Mismatches 191; Indels 108; Gaps 19;

QY 10 PVLVFGWLGIEISSGCEGLWLDARTCFVPMKFAKDLSEADARIFKAAVAVARG 69  
 DB 6 PRLI--PWLVSQDLGLQEGVAMWNKSRTRFRIPWKGHGLQDAQOEDGFQMAEATGA 63  
 QY 70 WPPSRGGCPPEAEATARAGMKTNRCAALSTRFVWLRLNSGDPADEPKVALSLREL 129  
 DB 64 YVP---GRDKPLPT---WKRNFRSALNRKRGRLADRSKDPDPKHYEVV---110  
 QY 130 WREGPGTDQTEAAAPAVPPPGGPPGFLAHTHAGLQAPGAPAPAGDKGDLLOAVQ 189  
 DB 111 -NSGVG---DFGQPDTSPTNG-----GSTSD-----TQ 136  
 QY 190 SCLADHLTLTASGADPVPTAPEGQGLPLTGACAGGRLPAGELYGMAVENTTPSPGQ 249  
 DB 137 EDIDELL--GNMVLAPLP-----DPGPP-----SLAVALPEPCPQ 169  
 QY 250 PALITTEAAPSPHQAEPYLSPPSPACTAVOEP--SPGALDVTIMYKRTVLQKVGH 308  
 DB 170 PLSPPSLDNPTP-----FNLGPSENPRLKLLVPGEEWFEVTAFFRGVQVFOQTISCP 223  
 QY 309 SCFLVYPPPAVATPDQOVAFPSPA--ELPDQOLRTYTELLLRHVAFGLHLELGPOLW 367  
 DB 224 EGRLLVQ--SEVGRITLPGWVITLPDPMSLTDRGWSYVHVHSLCLOGGLALMRAQGMW 282  
 QY 368 ARMGKCKVYVEG-----GPPGSPSPACLLPNCOTPIPDVFPQGVLEVPARBRSGPR 418  
 DB 283 AORLGHCHTWAYASELLPNSGHPDGE-----VPXDKGQVFDLQPTVLDLTF 333  
 QY 419 ARORRGSPRTIYVFGQDLSAGRPKESLVLKLEPWLGRVHLEGTORGVSLSDDLDL 478  
 DB 334 EGSGR--SPRYALWFCVGSNPDQDPWTKRLVMKVVFTCLALVAVARVGGASSLENT-V 391



QY 479 DCLISSANSU---YDIECFLEML 499  
 DB 392 DLHISNSHPLINSQYKAYLODI 415

RESULT 6  
 ID IRF3\_MOUSE STANDARD; PRT; 497 AA.  
 AC P56477;  
 DT 15-JUN-1998 (Rel. 36, Created)  
 DT 15-JUN-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Interferon regulatory factor 5 (IRF-5).  
 GN IRF5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_taxid=10090;

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lymph node;  
 RA Grossman A.; Kondo S.; Antonio L.; Mak T.W.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 DR EMBL; AF028725; AAB81997.1; -.  
 DR HSSP; P23906; 1IRF.  
 DR MGD; MGI:1350924; 1rf5.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERREGFCT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 KM Transcription regulation; DNA-binding; Nuclear protein.  
 DNA BIND 16 118 TRYPTOPHAN PENTAD REPEAT.  
 DOMAIN 9 12 POLY-PRO.  
 SQ SEQUENCE 497 AA; 56005 MW; D8BD54DB946E354F CRC64;

Query Match 16.3%; Score 444.5; DB 1; Length 497;  
 Best Local Similarity 29.2%; Pred. No. 8,7e-19;  
 Matches 148; Conservative 58; Mismatches 188; Indels 113; Gaps 21;

QY 11 RVLFGWLVGEISSGCEGLQWIDEARTCPVPWPKHFAKDLG--EADARIFKAVANAVARG 69  
 DB 14 RVRLKPLVAVQVNSCOYPLQWVNGEKKLFYIPRHATRHQPSODGQNTIFKAWKETGK 73  
 QY 70 WPPSSRGCGPPPEAETERAGWKTNFCALSTRFVMLRDNSSD--PADPHKYVALREL 128  
 DB 74 YTEG-----VDEADPAKMKNLKALSKSHDFOLFYDGPDMPPQPKYK-----EV 120  
 QY 129 CWRGCP-----GTDOTEAEAPAAVPPGCGPPGFPAHNAHAGLQAPGAPAP 175  
 DB 121 C-SNGCPAPTESQPTDDVVLGSEEEEBEL-----QMLRQ--LSTIRPAL--RGPVNPAP 170  
 QY 176 AGDXGDLILQAVQSCILADHLLTASWGADPVTPKAPGEGGLPLTGACAGGGLPAGEL 235  
 DB 171 -----YSLPKEDTKMPALQPPVGL-----GPPVDPNLT 199  
 QY 236 YGMAVETTPSPQO-----PALTTGGAAPBESPHQAEPTL-----SPSSACTAQAOP 284  
 DB 200 L-----APPSGNPAGFRQLPLPEVLEPGLASSQPP--TEPLLPDLLISPHMLPLT----- 247

QY 285 SPGALDVTIMYKGRVLYQKVGHP--SCTELYGPPDPAVATDPO-----QVAPPS 333  
 DB 248 ---DLRIKFOYGRABRTLTLSNPQGRLFYS---QLEATQOQVLEFGPVTLLEQVRFPS 300

QY 334 PALPPOKQRLTYTEELLRIHVARGLHLELRGPOLMARMKCKKYWVGVPSPGASSTPA 393  
 DB 301 PEDIPSDKORFYTNQLLDVLDRLGLITQLQGQDYALRLCQCKVW--SGPCALHSGCPN 358

QY 394 CLLPKNCDDPIPFDFRVPQELVEFRARQRRSPRYTYLGFQGLDSAGRPREKSLVLYKL 453  
 DB 359 -PIQREVKRLKLSLEPFLNELLIPOKQGNTPPPPFIFCFGEWMDVAVPREKLLTVQV 417

QY 454 EPMVLCRVHLEGTQREGVSSLDSSDDL 480  
 DB 418 VPVAARLLLEMFSGELMSADSIRLQI 444

RESULT 7  
 ID IRF3\_MOUSE STANDARD; PRT; 419 AA.  
 AC P70671;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Interferon regulatory factor 3 (IRF-3).  
 GN IRF3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_taxid=10090;

RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RA Hakem R.; Grossman A.; Antonio L.; Suggs S.; Mak T.W.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRO-1 (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 DR EMBL; U75839; AAB36924.1; -.  
 DR EMBL; U75840; AAB36925.1; -.  
 DR HSSP; P23906; 2IRF.  
 DR MGD; MGI:1859179; 1rf3.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF.1.  
 DR PRINTS; PR00267; INTERREGFCT.  
 DR PRODOM; PD002355; IRF.1.  
 DR SMART; SM00348; IRF.1.  
 DR PROSITE; PS00601; IRF.1.  
 KM Transcription regulation; DNA-binding; Nuclear protein; Activator.  
 FT DNA BIND 7 107 TRYPTOPHAN PENTAD REPEAT.  
 SQ SEQUENCE 419 AA; 46852 MW; 1FF67CAE0FC7F027 CRC64;

Query Match 16.0%; Score 438; DB 1; Length 419;  
 Best Local Similarity 27.6%; Pred. No. 1,7e-18;  
 Matches 143; Conservative 56; Mismatches 182; Indels 138; Gaps 17;

QY 6 ERAPVTLGFWLVGEISSGCEGLQWIDEARTCPVPWPKHFAKDLSEADARIFKAVAV 65  
 DB 2 ETRKPRIL--PWLVSQDILQGLEGVAMLDSSRRRFPIMWGHGRQDAQMAHDFGIPOKMAE 59  
 QY 66 ARGWPPSSRGCGPPPEAETARAGWKTNFCALSTRFVMLRDNSSGDPADPHKYVALS 125

```

Db 60 ASGAVTP-----GKXDPVST-----WKNFRSALNRKEVLRILAAANSKDPYDPKHVYEFV 110
Qy 126 RELCKREPGT-----DQTEAPAPAVPPPGGPPPPFLAHITAGLQAGPLPAPAGD 178
Db 111 T-----PGARDPVHLGASPTNGKSSLPHSQENLPFLF-----DGLILGPLKD 153
Qy 179 KGDILLQAVQOCLADHLLTASWGADVPPTKAPGEGGLPLTGACAGGPGLPAGELGW 238
Db 154 EGSSPLAIV-----SDP-----165
Qy 239 AVETTPSPGPQPAALTTEBAAPSPHOAEPLSPSPACTAVQSPSCAL-----DV 291
Db 166 -SQQLPSP-----NVNMFNPAPQ-----ENPLKQLAEQWMEVEV 200
Qy 292 TIMYGRVTLQKVHSPCTFLYGPDPDAVATDQQAFAFSP-ALPDQOLRYTEBL 350
Db 201 TAFYRGROVPQOTLCFPGGLRLVG--STADMTLPMQPVTLDPBEGFLTKLVKEVVGVL 258
Qy 351 RHVAPGLHLRLGPPOLMARMGCKVYWEVG-----GPPGASPSPTACILPRNCD 401
Db 259 KGLNGGLALMQAGQCLMAQRGLSHAFALGEBLLPDSGRGPDG-----VHKXD 309
Qy 402 TPIDFRVFPQELVEFRARQRGSPRTIYLGFGODLSAGRPKESLVLVKLEPWLGRVH 461
Db 310 GAVFLRPFVADLLIFMEGSGH-SPRYTLMFQMGEMFPQDPBWYKRLVMKVPCTLCEL 368
Qy 462 LEGTQREGVSSLDSDLDLCLSSANSL-YDIECTIMEL 499
Db 369 LEMAREGASSLKTVDLHINSQPISTSDQYKAYLOL 407

```

## RESULT 8

IRF6 MOUSE STANDARD: PRT: 467 AA.

ID IRF6 MOUSE STANDARD: PRT: 467 AA.  
AC P97431.  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Interferon regulatory factor 6 (IRF-6).

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCB1\_TaxID=10090;

RP SEQUENCE FROM N.A.  
RC STRAIN=BAUB/c; TISSUE=Colon;  
RA Grossman A., Mitrucker H.W., Antonio L., Mak T.W.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
DE -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
DE -1- SIMILARITY: BELONGS TO THE IRF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL; U73029; AAB36714.1; .  
DR HSSP; P23906; 2IRF.  
DR MGD; MGI:1859211; IrF6.  
DR InterPro; IPR001346; IRF.  
DR Pfam; PF00605; IRF; 1.  
DR PRINTS; PR00267; INTERNEGCT.  
DR PRODOM; PD002355; IRF; 1.  
DR SMART; SM00348; IRF; 1.  
DR PROSITE; PS00601; IRF; 1.  
KM Transcription regulation; DNA-binding; Nuclear protein.  
FT DNA BIND 9 111 TRYTOPHAN PENTAD REPEAT.  
SQ SEQUENCE 467 AA; 53106 MW; 68CCAA90680FEDC8 CRC64;

Query Match 16.0%; Score 436; DB 1; Length 467;  
Best Local Similarity 27.7%; Pred. No. 2,5e-18;  
Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 17;

```

Qy 1 MALAPERABRVLFGWLLGELISSGCEYGLQWIDEARCTFRVPWKIPAR-KDISEADARI 59
Db 1 MALTPR---RVRLKPLVAQVDSGLYFGLIMLHRSKRFQIPWKIATRHSPQEEBNTI 56
Qy 60 FKMAVAVRGWPPSSRCGGPPPEAETARAGMTNRCALRSRRFVMLRDNQGD-PADP 118
Db 57 FKMAAVETGKY---QEGVDDPDPK-----WRAQLKCALNKSREFLMTDGTKEVPMN 107
Qy 119 HKVYALSRELQWEGPGTQDTEAPAPAVPPPGGPPFLAHITAGLQAGPLPAPAGD 178
Db 108 VKIY---QYC-----DIQPGQ-----SVINPGTGSAPWD 135
Qy 179 KGDILLQAVQOCLADHLLTASWGADVPPTKAPGEGGLPL---TGACAGGGLPAGE 234
Db 136 EKQDVDEDEE---EDELQSOHNVPIQDTPFLININGSPPMAPASVGNCSVGNCSF--- 188
Qy 235 LYGNVAVETTPSPGPQPAALTTEBAAPSPHOAEPLSPSPACTAVQSPSCALDVTIM 294
Db 189 -----BSWPKTEPLEMEVPAQIOPFYSSPELWISSLPMT-----DLDTKFO 221
Qy 295 YKGRVTLQKV-VGHP-SCTFLYGPDP-----AVRATDQQAFAFSPAPLDPQOLRYT 346
Db 232 YRGVEGQMTVNSPQCRFLFYGDLGMPQOEBLFGVSLQYKFPPEHITNEKOKLFT 291
Qy 347 BELIRHVAAGLHLRLGPPOLMARMGCKVYWEVGPPGASPSPTACILPRNCDTPIFD 406
Db 292 SKLIDVMDRGLILEVSGHAIYAIRLCCKKYWMSGPCAPSLAASN-----LIEROKVVLFC 347
Qy 407 FRVFPQELVEFRARQRGSPRTIYLGFGODLSAGRPKESLVLVKLEPWLGRVHLEGTQ 466
Db 348 LETFLSELHMQQLEKQPPPEIYLCFGEWPDGKRLERKLLVVOIVVAVMYEMFS 407
Qy 467 REGVSSLDSDLDLCLSSANSLYDIECFIMEL 499
Db 408 GDFTRFSDSGSVRLQISTPD-INDNIVAQKQL 439

```

## RESULT 9

IRF6 HUMAN STANDARD: PRT: 467 AA.

ID IRF6 HUMAN STANDARD: PRT: 467 AA.  
AC O14856;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Interferon regulatory factor 6 (IRF-6).

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCB1\_TaxID=9606;

RP SEQUENCE FROM N.A.  
RC Grossman A., Mitrucker H.W., Antonio L., Ozato K., Mak T.W.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DE -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC  
RN [2]  
RN [3]

RP SEQUENCE FROM N.A.  
RC Grafham D.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RP TISSUE=Placenta;

RA Strauberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AF027292; AAB8411.1; -  
 DR EMBL; AL023298; CA18545.1; -  
 DR EMBL; BC014852; AAH14952.1; -  
 DR HSSP; P23906; 2IRF.  
 DR Genem; HGNC:6121; IRF6.  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF; 1.  
 DR PRINTS; PR00267; INTERFERECT.  
 DR ProDom; PD002355; IRF; 1.  
 DR SMART; SM00348; IRF; 1.  
 DR PROSITE; PS00601; IRF; 1.  
 DR Transcription regulation; DNA-binding; Nuclear protein.  
 DR DNA BIND 9 TRYPTOPHAN PENTAD REPEAT.  
 DR SEQUENCE 467 AA; 53129 MW; 7E28F5E0F5BA053 CRC64;

Query Match 15.8%; Score 431; DB 1; Length 467;  
 Best Local Similarity 27.7%; Pred. No. 4.8e-15;  
 Matches 142; Conservative 72; Mismatches 211; Indels 88; Gaps 20;

CC 1 MALAPEAARVVLGEWILGEGISGCEGLQWLDDEARTCFRVPWKHFAK-KDLEADARI 59  
 CC 1 MALHPR-----RVRLKPLVAVQVDSGLYGLILWHRDSKRFQIPKHAIRISPOEEENT 56  
 CC 60 FKAAVAVRGWPPSSRGSGPPPEAETAEKAGKTNFCALRSTRFVWLKDNQSD-PADP 118  
 CC 57 FKAAVAVETGKY--QEGVDDPDPK-----WKQLRCALINKSHFELMTDGTKEVPMNP 107  
 CC 119 HKYVALSRELCKMEGPTDQTEAARAAPVPPGCGPPGFLAHAGLQAPGLPAPACD 178  
 CC 108 VKYI-----QVC-----DIQPOG-----SINPGSTGAPWD 135  
 CC 179 KGLDLLQAVQOSCLADHLTASMGADVPPTKAPGEGEGPL-----TGACAGGPGLPAGE 234  
 CC 136 EKD---NDVDEDEDE-LOOSQHNVYIQTFFPLNINGSBMAASGNGSVNGCSBAV 191  
 CC 235 LYGVAVETTPSPGPPALTTGGAAPSPHQAEPVSPSPACTAOPSPGALDVTIM 294  
 CC 192 ---W-----PKTEPLEMEVPQ--ADIQPFYSSELMWISSLPMT-----DLIDIKFQ 231  
 CC 295 YKGRITVQKY-VGHP-SCTFLYGPDP-----AVRATDQOVAFPSPALPDOKQRYT 346  
 CC 232 YRGEKYGQITVSNPQGRFLFYGLDGPMPQOEBLFGVSLQYKFPGRPHITNEKQKFLT 291  
 CC 347 EELLRHVAPGLHLELRGPOLMARMGKCYVWEVGGPSPGASPTPACLLPRNCDTPIFD 406  
 CC 292 SKLIDVMDRLILEVSGHAIYAIHLCOCKYVWSPCAPSLVAPN-----LIERQKVVLCF 347  
 CC 407 FRVVFQELVEFRARQRGSPRYTYLFGQDLSAGRPKESLVLYKLEPWLCRYHLEGTQ 466  
 CC 348 LKTFPLSLIHAQKQIKQPPFEIYLCFGEEMWPGKLERLILVQYIPVAVARIYEMFS 407  
 CC 467 REGVSSLDSSDDLCLSSANSLYNDIECFMEL 499  
 CC 408 GDFTRSPDSSGVRQLQISTPD-IKDNIVAQLKQL 439

RESULT 10  
 IRF4\_MOUSE  
 ID IRF4\_MOUSE STANDARD; PRT; 450 AA.  
 AC Q64287; Q60802;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Interferon regulatory factor 4 (IRF-4) (Lymphocyte specific interferon  
 DE regulatory factor) (LIRF) (NF-E45) (PU.1 interaction partner)  
 DE (transcriptional activator PIP).  
 GN IRF4 OR SPIP.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=95317607; PubMed=7797077;  
 RA Eisenbeis C.F., Singh H., Storb U;  
 RT "p12, a novel IRF family member, is a lymphoid-specific,  
 RT PU.1-dependent transcriptional activator."  
 RL Genes Dev. 9:1377-1387 (1995).

[2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; and 129/SvJ; TISSUE=Spleen;  
 RX MEDLINE=9534364; PubMed=7541907;  
 RA Matsuyama T., Grossman A., Mittlemeier H.-W., Siderovski D.P.,  
 RA Kleier F., Kawakami T., Richardson C.D., Taniguchi T., Yoshinaga S.K.,  
 RA Mak T.W.;  
 RT "Molecular cloning of LsIRF, a lymphoid-specific member of the  
 RT interferon regulatory factor family that binds the interferon-  
 RT stimulated response element (ISRE)."  
 RL Nucleic Acids Res. 23:2127-2136 (1995).

CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-  
 CC STIMULATED RESPONSE ELEMENT (ISRE) OF THE MHC CLASS I PROMOTER.  
 CC BINDS THE IMMUNOGLOBULIN LAMBDA LIGHT CHAIN ENHANCER, TOGETHER  
 CC WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL  
 CC TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.

CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE  
 CC OR ABSENCE OF A GLUTAMIC ACID RESIDUE.

CC -1- TISSUE SPECIFICITY: LYMPHOID CELLS.  
 CC -1- INDUCTION: NOT INDUCED BY INTERFERONS.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.

CC -----  
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U34307; AAA75283.1; -  
 CC EMBL; U11692; AAA75309.1; -  
 CC EMBL; U20949; AAA75316.1; -  
 CC EMBL; U20949; AAA75317.1; -  
 CC HSSP; P23906; 2IRF.  
 CC MGD; MGI:1096873; Irfa.  
 CC InterPro; IPR001346; IRF.  
 CC Pfam; PF00605; IRF; 1.  
 CC PRINTS; PR00267; INTERFERECT.  
 CC ProDom; PD002355; IRF; 1.  
 CC SMART; SM00348; IRF; 1.  
 CC PROSITE; PS00601; IRF; 1.  
 CC Transcription regulation; DNA-binding; Nuclear protein; Activator;  
 CC Alternative splicing.  
 CC DNA BIND 23 TRYPTOPHAN PENTAD REPEAT.  
 CC FT VARSPLIC 165 MISSING (IN ISOFORM 2).  
 CC SEQUENCE 450 AA; 51577 MW; 5FD94CA6C453869C CRC64;

Query Match 13.8%; Score 376; DB 1; Length 450;  
 Best Local Similarity 28.1%; Pred. No. 6.3e-15;  
 Matches 130; Conservative 66; Mismatches 172; Indels 94; Gaps 22;

CC 16 EWLIGRISGCEYGLQWLDDEARTCFRVPWKHFAKDLSEADARIFKAAVAVRGWPPSS 74  
 CC 26 QWILIDQIDSKYGLVWENDEKSVFRLIPMKHAKQKYNEEDALFKAAVLFKGF----- 81  
 CC 75 RGGGPPPEAETAEKAGKTNFCALRSTRFVWLKDNQSD-PADPDKVIALSRELCKRGG 133  
 CC 82 REGIDRDPDPT-----WKTRRLCALINKSNDFELVRSQDISDPYKVVYRI----- 127

134 PGTDTAEAPAAVPPQGGPRP---FLAHTAGLQAGPLPAPAGKDLLLQAVOQS 190  
 128 -----VPGAKKGAQQLTLDTOAMGHPRPTAP---YGSIPAOQVANY 169  
 191 CLADHLTASWAGADVPYTKAPGEGOEGLPLTGACAG---GFGLP---AGELYGMAVE 241  
 170 MMPPH--DRSW-RDVAPOSHPEIRYQCGVTFGRGHMHQSGSCNGQVGTFAACAP 226  
 242 TTPSPG-PQPAALTTGEAAPSPHQAEPYLSPSACTAVOEPSGALDVTIMKGTIV 300  
 227 ESAPGPIEPSIRSAEALA-----LSDC-----RLHICLYRDLIV 263  
 301 LOKVGHF-SCFELYGPPDPAVRATDPQOVAFPSPAEIPDOKOLYTEELLRHVAPGLH 359  
 264 KELTTSPGCRISG---HTYDASNLDQVLPYP---DDNGQRNIKLSHLERGLV 317  
 360 ELRGQILWARMKCKVYWEVGGPPGAS--PSTPACLLPNCPTPIFDVFPQELVEF 417  
 318 WMAPDGLYAKRLCGSRITWD--GPLALCSDRPNK---LENDQCKLFDTOQFSELQVF 371  
 418 RARQRRGSPRTYIYFGQDLSAGPKKSLVVLVEWLCR 459  
 372 -AHGRRPAPRQVTLCFGEPEPPDPQ-RQRKLITAVEBLAR 411  
 RESULT 11  
 IRF4 HUMAN STANDARD; PRT; 451 AA.  
 AC Q15306; Q99660;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Interferon regulatory factor 4 (IRF-4) (lymphocyte specific interferon regulatory factor) (USIRF) (NF-EWS) (multiple myeloma oncogene 1).  
 GN IRF4 OR NM01.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Spleen;  
 RX MEDLINE=97079690; PubMed=8921401;  
 RA Grossman L., Mitrucker H.W., Nicholl J., Suzuki A., Chung S., Antonio L., Suga S., Sutherland G.R., Siderovski D.P., Mak T.W.; "Cloning of human lymphocyte-specific interferon regulatory factor (hUSIRF/hIRF4) and mapping of the gene to 6p23-p25.";  
 RT Genomics 37:229-233(1996).  
 RL [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Spleen;  
 RC TISSUE=Skin;  
 RA Strauberg R.;  
 RA Corradini P., Boccadoro M., Nilsson K., Chaganti R.S.K., Dalla-Pavera R.; "Deregulated expression of MDM1 gene by t(6;14)(p25;q32) chromosomal translocation in multiple myeloma.";  
 RT Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Skin;  
 RA Strauberg R.;  
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE INTERFERON-STIMULATED RESPONSE ELEMENT (ISRE) OF THE MFC CLASS 1 PROMOTER. BINDS THE IMMUNOGLOBULIN LAMBDA LIGHT CHAIN ENHANCER, TOGETHER WITH PU.1. PROBABLY PLAYS A ROLE IN ISRE-TARGETED SIGNAL TRANSDUCTION MECHANISMS SPECIFIC TO LYMPHOID CELLS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER IN THE PRESENCE OR ABSENCE OF A GLUTAMIC ACID RESIDUE.  
 CC -1- TISSUE SPECIFICITY: LYMPHOID CELLS.

CC -1- INDUCTION: NOT INDUCED BY INTERFERONS.  
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ch/announce/or send an email to license@ebi.ch>).  
 CC -----  
 DR EMBL; U52682; AAC50779.1; -  
 DR EMBL; U63738; AAB37258.1; -  
 DR EMBL; BC015752; AAH15752.1; -  
 DR HSSP; P23906; 2IRF.  
 DR TRANSFAC; T04929; -  
 DR GeneW; HGNC; 6119; IRF4.  
 DR MIM; 601900; -  
 DR InterPro; IPR001346; IRF.  
 DR Pfam; PF00605; IRF; 1.  
 DR PRINTS; PR00267; INTERFERGCT.  
 DR PRODOM; PD002355; IRF; 1.  
 DR SMART; SM00348; IRF; 1.  
 DR PROSITE; PS00601; IRF; 1.  
 KW Transcription regulation; DNA-binding; Nuclear protein; Activator;  
 KW Alternative splicing.  
 FT DNA BIND 23 125 TRYPTOPHAN PENTAD REPEAT.  
 FT VARSPIC 165 165 MISSING (IN ISOFORM 2).  
 FT CONFLICT 300 300 Q -> H (IN REF. 2).  
 FT CONFLICT 306 306 K -> N (IN REF. 2).  
 FT CONFLICT 333 333 R -> T (IN REF. 2).  
 SQ SEQUENCE 451 AA; 51772 MW; 17CD1327C6F5BFPA CRC64;  
 Query Match 13.3%; Score 363; DB 1; Length 451;  
 Best Local Similarity 27.5%; Pred. No. 3.5e-14;  
 Matches 136; Conservative 73; Mismatches 193; Indels 92; Gaps 24;  
 16 EMLIGEISGCEYLQWLDARTCFRVMGHFARKDLS-BADRIFKAMAVARGWPSS 74  
 26 QMLIDQIDSGKRYPLVVENEEKSIFRIPMGHAGKQDYNREDDALFQAMLFQKGF---- 81  
 75 RGGGPPPEABTAEBAKKTNFRCALSTRFVMLRDSNG--DPADPHKVALISBELCMBEG 133  
 82 REGIDKDDPT-----WKTRLRCALNKSNDFEELVERSQDISDPYKYRIIVEGAKK-- 134  
 134 PGTDTAEAPAAVPPQGGPRPFLAHTAGLQAGPLPAPAGKDLLLQAVOQSL 192  
 135 -GAKQLTLEDPQW-----SMSPYPT-MTTPYPSLPA-----QQVHNYVM 171  
 193 ADHLTASWAGADVPYTKAPGEGOEGLPLT-----GACAGGPGLPAGELYGMAVE 242  
 172 PP--LDRSW-RDVPDPHPEIRYQCGVTFGRGHMHQSGACNGCQV-TGTFACAPRE 227  
 243 TTPSPG-PQPAALTTGEAAPSPHQAEPYLSPSACTAVOEPSGALDVTIMKGTIV 301  
 228 SOAGVGTPEPSIRSAEALA-----FSDC-----RLHICLYRDLIV 264  
 302 QKVGHF-SCFELYGPPDPAVRATDPQOVAFPSPAEIPDOKOLYTEELLRHVAPGLH 360  
 265 ELTTSPEGCRISG---HTYDASNLDQVLPYP---EDNGQRNIKLSHLERGLV 318  
 361 LRGPQLWARMKCKVYWEVGGPPGAS--PSTPACLLPNCPTPIFDVFPQELVEF 418  
 319 WMAPDGLYAKRLCGSRITWD--GPLALCSDRPNK---LENDQCKLFDTOQFSELQAF- 371  
 419 ARQRRGSPRTYIYFGQDLSAGPKKSLVVLVEWLCRVLHEGTQREGVSSLDSSDL 478  
 372 AHGRRPAPRQVTLCFGEPEPPDPQ-RQRKLITAVEBLARQLRYVFAQNSGHLRGLYDL 430  
 479 DLCLSSANSLYDDI 492  
 431 PEHISNEDYHRSI 444

```

Db      186 PS-----LPLVNGTGVGEQHSHGVSQWVITFFYSGR----- 216
Qy      304 VGHPSCTFLYG-----PPDAVNRATDP---QOVAFPSPAEPLPDOKUITYBELLRHV 353
Db      217 LVGHITITVSEYEGCRSLSPSPNSHGEKLYTPEDSLSEHVFPFAEA1QNDRQOITKTLGHL 276
Qy      354 APGLHLELRPQOLMARBMGCKYVWEVGGPFGSASPSPTACLPRNCDTPIIDPRVFPQOE 413
Db      277 ERGLHLHSHKQGIIFKILCGGRFWSGNTVYVYDRPSK-----LDRDSVVKIFDTNLPFRE 332
Qy      414 LVFERARQRRGSPRYTITLGFQODLSAGRPKESKSLVYVLEKPLMCLRCHALEGTQREGVSSL 473
Db      333 LQGYVNNQGR-FPDSRYMLTFGESEFPPTVPLRCLILVQYEQLCVRYMGEAGKTCSSPM 391
Qy      474 DSSDL-----DIC 481
Db      392 LPDVOQEQYVRIFFODIC 409

RESULT 13
ICSB_MOUSE
ID      ICSB_MOUSE      STANDARD;      PRT;      424 AA.
AC      P23611;
DT      01-NOV-1991 (Rel. 20, Created)
DT      01-NOV-1991 (Rel. 20, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      Interferon consensus sequence binding protein (ICSBP).
GN      ICSBP1 OR ICSBP.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90251633; PubMed=2111015;
RA      Driggers P.H., Ernst D.L., Gleason S.L., Mak W.-H., Marks M.S.,
RA      Levi B.-Z., Flanagan J.R., Appella E., Ozato K.,
RT      "An interferon gamma-regulated protein that binds the interferon-
RT      inducible enhancer element of major histocompatibility complex class
RT      I genes."
RL      Proc. Natl. Acad. Sci. U.S.A. 87:3743-3747(1990).
CC      -I- FUNCTION. SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION
CC      OR TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON
CC      CONSENSUS SEQUENCE (ICS)). PLAYS A REGULATORY ROLE IN CELLS OF THE
CC      IMMUNE SYSTEM.
CC      -I- SUBCELLULAR LOCATION: Nuclear.
CC      -I- TISSUE SPECIFICITY: PREDOMINANTLY IN LYMPHOID TISSUES.
CC      -I- INDUCTION: BY INTERFERON GAMMA.
CC      -I- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
Db      EMBL; M32489; AAA37878.1; -
Db      PIR; A35861; A35861.
Db      HSSP; P23906; 21RF.
Db      TRANSFAC; T00402; -.
Db      MGD; MGI:96395; Icsbp.
Db      InterPro; IPR001346; IRF.
Db      Pfam; PF00605; IRF; 1.
Db      PRINTS; PR00267; INTERNREGFCT.
Db      ProDom; PD002355; IRF; 1.
Db      SMART; SM00348; IRF; 1.
Db      PROSITE; PS00601; IRF; 1.
KW      Transcription regulation; DNA-binding; Activator; Nuclear protein;
KW      Interferon induction
FT      DNA_BIND          9      110      TRPPTOPHAN PENIAD REPEAT.
FT      Q      424; AA; 48237 MW; FBEF79A76846E8B82 CRC64;

```



```

RESULT 15
ICSB_HUMAN STANDARD; PRT; 426 AA.
ID ICSB_HUMAN STANDARD; PRT; 426 AA.
AC 002556;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon consensus sequence binding protein (ICSBP).
GN ICSBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Lung, and Monocytes;
RC MEDLINE=93094284; PubMed=1460054;
Weisz A., Maix P., Sharf R., Appella E., Driggers P.H., Ozato K.,
Levi B.-Z.;
"Human interferon consensus sequence binding protein is a negative
RT regulator of enhancer elements common to interferon-inducible
RT genes.";
RL J. Biol. Chem. 267:25589-25596(1992).
RN [2]
RP REVISIONS.
RA Schmitt M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SPECIFICALLY BINDS TO THE UPSTREAM REGULATORY REGION
CC OF TYPE I IFN AND IFN-INDUCIBLE MHC CLASS I GENES (THE INTERFERON
CC CONSENSUS SEQUENCE (ICS)). PLAYS A NEGATIVE REGULATORY ROLE IN
CC CELLS OF THE IMMUNE SYSTEM.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN LYMPHOID TISSUES.
CC -!- INDUCTION: BY INTERFERON GAMMA.
CC -!- SIMILARITY: BELONGS TO THE IRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M91196; AAB63813.1; -.
PIR; A45064; A45064.
HSP; P23906; 2IRF.
TRANSPAC; T02038; -.
DR Genew; HGNC:5358; ICSBP1.
DR MIM; 601565; -.
DR InterPro; IPR001346; IRF.
DR Pfam; PF00605; IRF.1.
DR PRINTS; PR00267; INTERNBEGCT.
DR ProDom; PD002355; IRF.1.
DR SMART; SM00348; IRF.1.
DR PROSITE; PS00601; IRF.1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Interferon induction.
FT DNA BIND 9 110 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 426 AA; 48356 MW; 1535DB7C83E0355 CRC64;
Query Match 11.3%; Score 308.5; DB 1; Length 426;
Best Local Similarity 21.8%; Pred. No. 4.2e-11;
Matches 116; Conservative 71; Mismatches 161; Indels 185; Gaps 18;
QY 16 EWLIGEISGGYEGIQIDERTCFRVWKFARKDLS-EDARIFKAMAVARGRPSS 74
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
12 QWLIEQIDSSMYPGIWIENEKSMFRIFPKAGKQDYNOQVDASIFKAMAVFKGKF- 67
QY 75 RGGGPPPEAETAEARAGKTNFRCLRSRRFVMLRDNSG-DRADPHKYVALSRELQREG 133
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 -----KEDGKAEPATWTKRLCALNKSPPDEEVTRSDLDISEPKYVRI----- 112

```

```

QY 134 PGTDQTEAEAPAAVPPQGGPFGFLAHTHAGLQAPRLPAPADKDDLILQAVQSC-- 191
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 --VPEEQKCKLGV-----ATAG-----CVNEVTMEGCR 140
QY 192 -LDHLLTASMGADPVPTKAPGEGQELPLTGACAGGPGLPAGLXGMAYETTSPPQRP 250
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 SEIDELIKE-----PSVDYMGMTKRSPP- 165
QY 251 AALTTGEAAPESPQAEPLSPSPSACTA-----VQSPSPA----- 288
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 -----PEKCRQLBDMWMAQPSCTVPLVTGTTTDAHSA 201
QY 289 --LDVTIMKGRVTLQKVVGP-SC-----TLYGPPDPAAVRATDPQVAFPS 333
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 FSGWVIFSYFGKLVGQATTTCPEGCRSLSPQGLPFTKLXGP-----EGLELVAFPP 254
QY 334 PAELPDQKQRYTEELLRHVAPGLHLELRPQLMARMGCKYWEVGGEPGSSPSTPA 393
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 ADAIPSERQQRQVTRKLFGLHREGVLLHSSRQGVFVKRLCGRFVCSGNAVVCCKGRPNK-- 312
QY 394 CLPRNCDTPIPFPRVFPQELVEFRARQRRGSPRYTYILGFGODLSAGRPKESLYLVKL 453
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 --LERDEVQVFTDSQFFRELQFPYNSQGR-LPDGRVVLCPGEFFPMAPLRSLTILVQI 369
QY 454 EPMICRVHLEGTORQEGVSSLDSSDL-----DICTSSANSLSLY 489
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 E-----QLYVQLAEAGKSCGAGSVMAQPEPPPDQVFRMPDICASHQSRSF 418

```

Search completed: June 18, 2003, 12:47:06  
Job time : 17 secs

